Efficient Estimation of Risk Measurement via Regression and Stochastic Mesh Method

XIONG, Ying

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Philosophy

in

Systems Engineering and Engineering Management

The Chinese University of Hong Kong

September 2011

Thesis/Assessment Committee

Professor Youhua Chen (Chair)

Professor Nan Chen (Thesis Supervisor)

Professor Kwai Sun Leung(Committee Member)

ABSTRACT

In recent years, portfolio risk measurement has turned out to be not only interesting and appealing but also very meaningful topic. In this thesis, we consider the problem of estimating the risk measurement for portfolio by nested simulation procedures. Firstly, we apply two methods developed by Gordy and Juneja (2006, 2008) and Broadie, Du, and Moallemi (2010). In their methods, Gordy and Juneja propose a standard nested simulation; Broadie, Du, and Moallemi present another efficient approach by sequentially allocating the computational budgets. Based on their experience, we develop two new approaches - Least Square Monte-Carlo approach and Stochastic Mesh approach in risk measurement. We use these two methods to estimate the future portfolio value as a conditional expectation in the inner level simulation step. Moreover, several numerical experiments have been conducted and we compare the four approaches numerically and find out that Least Square method may be fast and easy to implement without allocating computational budgets into outer and inner level simulations, while stochastic mesh method can be accurate but time-consuming.

Key words: nested simulation, loss distribution, value-at-risk, expected shortfall, stochastic mesh, least square approach.

摘要

近年来投资组合的风险管理问题成为一个十分吸引并且深具现实意义的课题。在这篇论文中，我们主要考虑用嵌套模拟程序估计投资组合的风险测量的问题。我们首先运用了Gordy和Juneja (2006, 2008)以及Broadie, Du, 和Moallemi (2010)提出过的两个方法。Gordy和Juneja主要提出了一种最标准的嵌套模拟的算法，而Broadie, Du, 和Moallemi在此方法上改进出一种更有效率的算法，它是有次序的分配计算预算量。在此经验上，我们提出了两种新算法 - 最小二乘法和随机网格法。 我们用这两种算法来估计投资组合在将来某一时刻的价值，也就是在内层模拟时的一个条件期望值。除此以外我们进行一系列的数值实验来比较这四种算法。我们发现最小二乘法不用考虑如何分配内外层的计算预算，因而是快速并较易实现的。 而随机网格法虽然可以较准确给出估算值，但耗时较久。

关键字: 嵌套模拟，亏损值分布，风险价值，期望损失，随机网格，最小二乘法

Acknowledgements

I would like to express my greatest gratitude to my supervisor, Professor Nan Chen, for his patient guidance in development of the idea in this research as well as in writing this thesis.

Many thanks are given to Professor Kwai Sun, Leung and Professor Youhua, Chen for serving as members of my thesis committee.

Moreover, I would like to thank my friend Xiangwei, Wan who discussed with me about research and ideas, and I also want to thank my fellow members and friends in Chinese University of Hong Kong.

Finally, and most importantly, I would like to thank RGC GRF CUHK411110 for partly supporting this Mphil thesis.

Contents

List of Figures

CHAPTER 1

INTRODUCTION

1.1. Background and Objective

In 2008, one of the most shocking financial events was the bankruptcy of Lehman Brothers. The immediate trigger, of course, was the subprime mortgage crisis which brought a series of terrible impacts on global financial market. However, the firm's lack of attention to risk might take a role inevitably. In fact, the measurement and management of risk is vital to the survival of financial institutions and the stability of the financial system. Regulation requires each financial institution to measure the risk of the firm's entire portfolio and to set its capital reserves accordingly, also to reduce the chance of bankruptcy if large losses occur. To maintain the stability of financial industry, the Basel Committee on Banking Supervision created an international standard about how much capital banks needed to put aside to guard against financial and operational risks they face. It is known as the Basel Accord (Basel I, 1988 and Basel II, 2004). These capital requirements against possible losses are crucial aspects of regulatory efforts to prevent a cascade of defaults that can paralyze the global financial system. That's why risk measurement have turned to be not only interesting but also very meaningful topic, which my research work would like to focus on.

In this thesis, we attempt to efficiently estimate three different risk measurements - Probability of large loss, Value-at-Risk and Expected Shortfall based on

Monte-Carlo simulation methods. At first we review the two past methods in nested simulation for risk measurements. Then we apply the idea of Least Square Monte-Carlo and Stochastic Method in evaluating the conditional expectation values in nested simulation. Following the algorithms, we develop the numerical experiments to compare the convergence rate of the four methods under the one-dimensional state variable case and multi-dimensional case.

1.1.1. Risk Measurement

While, risk, in simple terms, measures how volatile a portfolio's returns are. Taking it as a probabilistic or statistical concept, there are a lot of (sometimes even conflicting) notions and measurements of risk. As a result, it can be difficult to measure the risk of a portfolio and determine how various investments and asset allocations affect the risk. So it is important to express the risk in a way that permits it to be understood and controlled by different people like traders, senior managers, regulators, and others. Traders need to have better understanding of how risk arises and, more importantly, how it can be averted so that their performance can be enhanced. Regulators and risk managers need to have certain tools to control risk and even to exploit it.

Value-at-Risk

In the past few decades, there were plenty of risk measures: historical standard deviations to measure portfolio's volatility, gap analysis to give a crude idea of interest-rate risk exposure, various duration concepts for fixed income assets, Greek parameters of derivatives, etc. Among these risk measurements, Value-at-Risk (VaR) is particularly appealing and valuable in portfolio management applications.

VaR offers a way to meaningfully aggregate different types of market risk with a statistical framework and therefore helps in measuring and communicating risk information. Such concept was nicely described by Linsmeier and Pearson

(Linsmeier and Pearson,1996,p.3). Philippe Jorion, in a milestone book (Value-at-Risk, 2001) on VaR, defined the VaR this way: "VaR summarizes the expected maximum loss (or worst loss) over a target horizon within a given confidence level". For example, if a bank's 10-day 99% VaR is \$3 million, there is only a 1% chance that losses will exceed \$3 million in 10 days. Since it is based on probabilities, it cannot be relied on with certainty, but is rather a level of confidence, which is selected by the user in advance.

VaR is recommended in the Basel II framework as the basis for capital requirements against market risk. Variance-covariance method, Historical simulation and Monte-Carlo simulation are the three main methods for computing VaR.

Expected Shortfall

VaR is often criticized as not presenting a full picture of the risks a company faces. Artzner et al.(1997,1999) have proposed the following two shortcomings of VaR:

- VaR measures percentiles of portfolio loss distributions, and discards extreme loss beyond the VaR level. Thus, VaR may ignore important information regarding the tails of the distribution. The BIS (Bank for International Settlements) Committee on the Global Financial System (2000) terms this problem as "tail risk".

- The second shortcoming is that VaR is not coherent. It does not satisfy the subadditivity property.

To remedy the problems, Artzner et al. (1997) proposed the use of expected shortfall which was defined as the conditional expectation of loss given the loss is beyond the VaR level. In other words, where VaR answers the question "how bad can things get?", expected shortfall answers "if things do get bad, what is our expected loss?". By its definition, expected shortfall took losses beyond the VaR level into account. It was demonstrated the expected shortfall had no tail risk under more lenient conditions and to be subadditive (Yamai and Yoshiba, 2002c).

Expected shortfall, like VaR, is a function of two parameters: the time horizon(N-day) and the confidence level($x\%$). It is the expected loss during an N-day period, conditioning on that the loss is greater than the xth percentile of the loss distribution. For example, with $x = 95$ and $N = 10$, the expected shortfall is the average amount which is lost over a 10-day period, assuming that the loss is greater than the 95th percentile of the loss distribution.

Computing Method based on Simulation

Simulation methods are often used to compute risk measurements. We briefly introduce historical simulation method and Monte-Carlo simulation as follows.

For the historical simulation, the idea is intuitively simple to understand. For instance, to calculate VaR, we need simply keep a historical record of daily profit and losses within the portfolio and then calculate the fifth percentile for 95 percent or one percentile for 99percent VaR. As well as being simple, the historical method is more realistic. Unlike variance-covariance method to calculate VaR, the normally distributed return assumption and constant correlations assumption are not needed in this case. On the other hand, the weakness is that only past information is included. For historical simulation, we use past data to predict the future performance of the portfolio losses. Therefore, the prediction is not guaranteed to be reliable most of the time.

Unlike historical simulation, the Monte-Carlo method does not conduct the simulation using the observed data over the last N periods to generate N hypothetical portfolio losses(or profit). Instead, one chooses a statistical distribution that is believed to adequately capture or approximate the possible changes in the market risk factors. Then thousands, or perhaps tens of thousands randomly generated simulations are run forward in time. Finally, the distribution of possible portfolio losses may be constructed. And the VaR is determined from the distribution. This method is more realistic and more likely to give an accurate result.

1.1.2. Monte-Carlo Simulation

The Monte-Carlo method has a long history in science and engineering, and it was first developed in the 1940s by John von Neumann, Stanislaw Ulam and Nicholas Metropolis to deal with some of the calculations involved in nuclear physics. It hasn't been used in the finance area until the late 1970s to price derivatives and estimate their Greeks. Glasserman (2004) developed the use of Monte Carlo method in finance and also used simulation as a vehicle for presenting models and ideas from financial engineering. The basic concept of Monte-Carlo method applied in financial engineering is to simulate repeatedly a random process for the financial variable of interest covering a wide range of possible situations. Thus simulation can create the entire distribution of portfolio values. In that case, Monte-Carlo simulation is more flexible and accurate.

The simulation process may involve some typical steps. At first we select a model for the stochastic variables of interest. Having chosen the model, we give an estimation to its correlated parameters - volatilities, correlations and so on based on the suitable judgements (historical or available market data). We then construct simulated paths for the stochastic variables. For each path, a set of hypothetical terminal values in the portfolio is produced. Then we repeat the simulation enough time to be confident that the simulated distribution of the portfolio value is sufficiently close to the real distribution. At last, we can obtain the goal estimate upon this proxy distribution.

A nested simulation is often conducted to estimate the portfolio risk. The basic idea is described below. In an outer simulation, we generate several financial scenarios and for each outer scenario we estimate the portfolio values in an inner simulation progress. The process is first to simulate repeatedly a random process for the financial variable of interest covering a wide range of possible situations. Thus simulation can create the entire distribution of portfolio values. To be more flexible and accurate than other methods, Monte-Carlo simulation chooses a statistical distribution that is believed to adequately capture or approximate the possible changes in the market risk factors. Then thousands, or perhaps tens of thousands randomly generated simulations are run forward in time. Finally, the distribution of possible portfolio losses may be constructed. And the corresponding risk measurements are determined from the distribution.

As the most powerful and flexible method, Monte-Carlo technique has its main advantages: Firstly, the large numbers of scenarios generated provide a more reliable and comprehensive measure of risk than analytic models. Secondly it can capture convexity of nonlinear instruments and changes in volatility and time. But the main obstacle here is also very obvious. Because of its high flexibility, Monte-Carlo method can potentially account for a wide range of risk; meanwhile for a meaningful result, the number of simulations can often increase rapidly to millions, which makes it highly computational-intensive.

1.2. Literature Review

Since the late 1960s, the phenomenal growth in trading activities and the massive increases in the range of instruments traded in the market have spurred the development of risk management. Regulatory agencies and investors are increasingly concerned about the risk exposures of financial institutions. Thus, risk management has been attracting tremendous interest in recent years and regarded as a distinct subfield of the theory of finance. Crouhy, Galai and Mark (2000) represented the overviews and consolidation of the entire financial risk management field including policies, methodologies, data, and technological infrastructure. McNeil, Frey and Embrechts (2006) provided a comprehensive treatment of the theoretical concepts and modeling techniques of quantitative risk management, as well as practical tools to solve real-world problems.

While managing the risk, the first step is to quantify the riskiness of our position and hence to decide if it is acceptable or not. For this reason, several classes of risk measures were proposed in literature. Among different kinds of risk measures, Value at Risk (VaR) and coherent risk measures were often considered. VaR was first used by major financial institutions in the late 1980s to measure the risk of their trading portfolios. Since that time period, the use of VaR has exploded. Linsmeier and Pearson (1996) and Duffie and Pan (1997) gave an overview for VaR, as well as the measuring methods.

Most risk practitioners embraced VaR with varying degrees of enthusiasm, but there were also those who warned that VaR incurred deeper problems and could be dangerous. As we know, it seems natural to look for a measure of risk which is "sensitive" to diversification. Unfortunately, in general VaR fails to satisfy this property and, even for sums of independent risky positions, its behavior is not as we would expect. Embrechts (2000) presented an overview of this criticism about VaR not being sub-additive, which meant that the risk of a portfolio could be larger than the sum of the stand-alone risks of its components when measured by VaR. (Also see, Artzner et al. (1997, 1999); Rootzén and Klüppelberg (1999), or Acerbi et al. (2001)) Hence, managing risk by VaR might fail to stimulate diversification. Moreover, VaR does not take into account the severity of an incurred damage event. As a response to these deficiencies, the notion of coherent risk measures was introduced in Artzner et al. (1997, 1999). An important example for a risk measure of this kind was the Expected Shortfall (Artzner et al., (2001)). Expected Shortfall is defined as the conditional expectation of loss beyond the VaR level. Acerbi and Tasche (2001) presented the coherence properties and some further background on Expected Shortfall.

In this thesis, nested simulation is involved to measure risk. We generate many possible future scenarios and try to value the portfolio value for each scenario. If derivative securities are involved, we may conduct an inner-level simulation to estimate the portfolio values. The resulting computational burden can be quite large. Based on this concern, there are mainly two approaches developed to make nested simulation more computationally efficient. The first approach referred to Frye (1998) and Shaw (1998), who proposed a dimension reduction method and interpolation. The main idea is first to perform Monte-

Carlo simulation to estimate the portfolio value in only some scenarios, then to use interpolation to approximate the portfolio values in all scenarios. The other approach is more automated and generic. The earliest work is the thesis of Lee (1998), which related to point estimation of VaR. Lee discussed how to reduce the mean squared error of the point estimator with a uniform nested simulation, which employed a constant number of inner samples across the portfolio re-valuation step, thus allocating computational effort uniformly across all scenarios. Then Gordy and Juneja (2006, 2008) followed the similar idea in proposing a simulation procedure for point estimation of a portfolio's large loss probabilities and Value-at-Risk. They analyzed how a fixed computational budget may be allocated to the inner and the outer step to minimize the mean square error of the resultant estimator. Broadie, Du, and Moallemi (2010) developed another efficient simulation procedure for estimating the probability of large loss by sequentially allocate computational effort in the inner simulation based on marginal changes in the risk estimator in each outer scenario. While, a two-level simulation for interval estimation of expected shortfall was presented by Lan, Nelson and Staum (2007, 2010). Liu and Staum (2010) also focused on expected shortfall and they provide a computationally efficient simulation procedure for point estimation of expected shortfall.

In this thesis, we develop two approaches to estimate the conditional expectation in portfolio repricing step - Least Square Approach and Stochastic Mesh Method. These methods were also used in American option pricing in Longstaff and Schwartz (2001) and Broadie and Glasserman (2004) respectively.

1.3. Structure of This Thesis

The rest of the thesis is organized as follows. In Chapter 2, we present the problem formulation and basic setting first, followed by the specific formula for different risk measurements under our problem setting. Then we will review the former two methods in nested simulation for risk measurement in Chapter

3. Section 3.1 covers the most common method - Uniform Sampling and its corresponding MSE estimator for different risk measures. In Section 3.2, the motivation and basic idea for Sequential Sampling have been introduced. In Chapter 4 we mainly discuss our two new approaches by applying n nested simulation for risk measurement. The algorithms are displayed in detail. There are some numerical experiments by using Monte-Carlo simulation showed in Chapter 5. At last, in Chapter 6 we give the insight into the future work and conclusion for the whole thesis.

CHAPTER 2

Problem Formulation and Review of Past Methods

2.1. Problem Formulation and Basic Setting

To describe the portfolio measurement problem in mathematical language, we first introduce the following notations. Let's consider the basic risk measurement in portfolio management at some future time $t = H$(the risk horizon).

Let S_t be a D-dimensional Markov process of several state variables that govern portfolio prices. The risk factor vector S_t might include equity prices and other underlying prices referenced by derivatives. We denote $\{\mathcal{F}_t\}$ as the filtration generated by S_t.

Consider the portfolio P consisting of K positions. Let T_k be the maturity of position k and assume T_k is finite. Let $C^{(k)}(t)$ be the cumulative cashflows for position k over the period $(0, t]$ which is based on the underlying S_t. So the market value of each position $V^{(k)}(t)$ is the discounted expected present value of its cash flows under the risk-neutral measure Q. We denote r as the riskless rate. Then

$$V^{(k)}(t) = E^Q[\int_t^{T_k} e^{-r(s-t)} dC^{(k)}(s)|\mathcal{F}_t] \quad for \;\; t < T_k$$

and

$$V^{(k)}(t) = 0 \quad for \;\; t \geq T_k$$

The total portfolio loss L can be computed by

$$L = \sum_{k=1}^{K}(V^{(k)}(0) - e^{-rH}(V^{(k)}(H) + \int_0^H e^{r(H-t)}\,dC^{(k)}(t)))$$

To explain the above equation, the first term in the above equation is the current value of the whole portfolio. The middle one is the discounted future value. At last we adjust for the interim cashflows which can be reinvested in the money market.

Without loss of generality, to simplify the notations and discuss different simulation approaches under a consistent assumption we assume the portfolio has no intermediate cashflows before time H, and the riskless rate is deterministic and during the period $t = (0, H]$ riskfree rate equals to 0 . Thus we can rewrite the expression for the portfolio loss.

$$\begin{aligned} L &= \sum_{k=1}^{K} V^{(k)}(0) - \sum_{k=1}^{K} V^{(k)}(H) \\ &\equiv V_0 - V_H \end{aligned} \tag{2.1}$$

where

$$V_H = \sum_{k=1}^{K} V^{(k)}(H) = E^Q(\sum_{k=1}^{K}\int_H^{T_k} e^{-r(s-H)}dC^{(k)}(s)|\mathcal{F}_t). \tag{2.2}$$

Here we assume the initial price V_0 is already known and can be taken as constants in our algorithm.

2.2. Risk Measurement

In mathematical language, risk can be defined as a functional ρ that quantifies the risk of the random variable L by a scalar $\rho(L) \in R$. In the introduction, we have already introduced the basic concepts for several risk measurement. Here, we will focus on the three most popular and widely used risk measurements: The probability of large loss, Value-at-Risk(VaR) and Expected Shortfall(ES). Here I add some more rigorous descriptions.

Probability of large loss The probability of large loss is perhaps the most basic risk measure. It is very straightforward and easy to understand. It tells a probability that the future portfolio value falls below a pre-specified threshold. We now give a mathematical definition:

For a given large loss u, the probability of a large loss is estimating the probability of portfolio loss L exceeding u, that is,

$$\alpha \equiv P(L > u).$$

Value-at-Risk Value-at-risk offers a way to meaningfully aggregate different types of market risk with a statistical framework and therefore helps in measuring and communicating risk information. It is defined as the loss level that will not be exceeded with a certain confidence level during a certain period of time.

To explain VaR more specifically, given a probability of α, or a confidence level $1-\alpha$, and a holding period of length t, then the $1-\alpha$ confidence VaR is the expected loss that will be less than the VaR with probability $1-\alpha$.

The mathematical expression for VaR is as below.

For a target probability α, VaR is the value l_α given by

$$l_\alpha = VaR_\alpha = \inf\{l : P(L \leq l) \geq 1-\alpha\}$$

Expected Shortfall As an alternative to VaR, Expected shortfall is a coherent measure of financial portfolio risk. This is also sometimes referred to as conditional value at risk(CVaR), average value at risk(AVaR), and expected tail loss(ETL).

Given a quantile-level α, ES is defined to be the expected loss of portfolio value given that a loss is occurring at or below the $\alpha - quantile$.

Thus we give the following formal definition for ES. For a given probability α, l_α is the corresponding VaR,

$$ES_\alpha = \frac{1}{\alpha}(E[L \cdot 1_{\{L \geq l_\alpha\}}] + l_\alpha(1 - \alpha - Pr(L < l_\alpha)))$$

where the first term is also known as "tail conditional expectation" and the second term is a correction for mass at quantile l_α.

If the underly distribution for L is a continuous distribution then the expected shortfall is equivalent to:

$$ES_\alpha = \frac{1}{\alpha}(E[L \cdot 1_{\{L \geq l_\alpha\}}])$$

In our thesis, we try to use monte-carlo methods to give estimations for these risk measurements. Suppose now we simulate n scenarios, $\omega_1, \ldots, \omega_n$ for the risk factors which the portfolio depends on and based on each scenario we obtain the portfolio loss $Loss(\omega_i)$ for $i = 1, \ldots, n$. Thus we approximate the different risk measurements as following.

The Probability of Large Loss For a given large loss u, the probability of a large loss can be estimated by,

$$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} 1_{\{\hat{L}(\omega_i) \geq u\}} \tag{2.3}$$

Value-at-Risk For a target insolvency probability α, to give an estimate for VaR_α, we sort these estimates $\hat{L}(\omega_1), \ldots, \hat{L}(\omega_n)$ as $\hat{L}_{[1]} \geq \hat{L}_{[2]} \geq \cdots \geq \hat{L}_{[n]}$. Then $\hat{L}_{\lceil \alpha n \rceil}$ provides an estimate of l_α.

$$\hat{l}_\alpha = \hat{L}_{\lceil \alpha n \rceil} \tag{2.4}$$

where $\lceil x \rceil$ is the ceiling of the real number x.

Expected Shortfall For a given probability α, l_α is the corresponding VaR, we already know the expected shortfall is

$$ES_\alpha[L] = \frac{1}{\alpha}(E[L \cdot 1_{\{L \geq l_\alpha\}}])$$

Thus

$$ES_\alpha[\hat{L}] = \frac{1}{\alpha}(E[\hat{L} \cdot 1_{\{\hat{L} \geq \hat{l}_\alpha\}}])$$

is the estimate for the expected shortfall.

we may write the point estimate as

$$\hat{ES}_\alpha = \frac{1}{\alpha} \sum_{i=1}^{\lceil n\alpha \rceil} \frac{1}{n} \hat{L}_{[i]} \tag{2.5}$$

where $\hat{L}_{[1]} \geq \hat{L}_{[2]} \geq \cdots \geq \hat{L}_{[n]}$

Two-level Simulation When we do risk measurement, Monte-Carlo simulation is a powerful tool. Basically, with Monte-Carlo technique, we choose the stochastic process and the probability distribution that generate time series of interest and then we create an incredibly high number of scenarios to evaluate the profit or losses on the portfolio of interest. For example, to estimate the probability of large loss of a specific portfolio, we firstly generate n scenarios: $\omega_1, \ldots, \omega_n$ for the risk factors which the portfolio depends on, then calculate the portfolio loss $Loss(\omega_i)$ for each scenario if the loss can be calculated precisely. At last given a loss threshold u, the probability of large loss is estimated via formula (2.3). This is a general single-level Monte-Carlo simulation. Unfortunately, the portfolio loss usually can not be exactly computed. That's the reason we need to develop a two-level simulation.

To obtain an estimate of the risk measurements related to portfolio loss, we need to approximate the distribution of the random variable $Loss$ by using Monte-Carlo. That is what we call an outer level simulation in which we draw simulation paths of sampling of risk factors up to the horizon H. Another challenge is that for each outer path of scenario generated, the exact value of portfolio loss is difficult to calculate. The portfolio often contains complex derivative securities which are path-dependent with nonlinear payoffs. As a result, an inner simulation procedure is applied to evaluate the portfolio value and cashflows between H and time maturities.

Since nested Monte-Carlo simulation can represent a heavy computational burden, there are various estimation approaches have been applied. I will mainly introduce the work of *Gordy and Juneja (2008)* and *Broadie, Du and Moallemi (2010)*.

2.3. Uniform Sampling

Gordy and Juneja (2008) developed a nested simulation for three risk measures. It follows the most common idea below. At first step, we simulate n outer sample paths. For each path, we uniformly distribute m inner samples for re-pricing the portfolio loss given information obtained upon some intermediate time H.

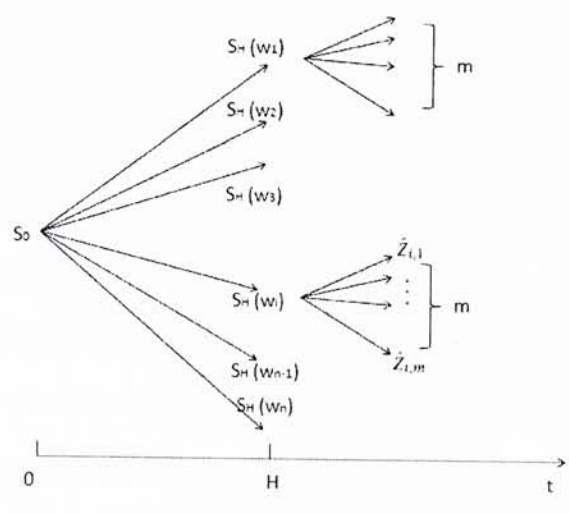

Figure 2.1: Uniform sampling

The simulation procedures are illustrated as follows.

- Draw n outer simulation paths S_t for $t \in (0, H]$ under the physical measure. Denote these n scenarios by $\omega_1, \ldots, \omega_n$.
- For each outer path, simulate m "inner step" trials from time H to the maturity. These inner trails are simulated under the risk-neutral measure.

- Evaluate portfolio price at maturity and discount back to time H for each inner path under the specific outer scenario to obtain the "V_H".
- Calculate portfolio loss L at time H using formula (2.1) for each inner path under the specific outer scenario. For "ω_i" scenario, denote $\hat{Z}_{i,1}, \hat{Z}_{i,2}, \ldots, \hat{Z}_{i,m}$ as m i.i.d. samples of losses generated according to the inner steps.

In the uniform sampling, there are uniformly m inner trails for each outer scenario ω_i. We can approximate $L(\omega_i)$ by

$$\hat{L}(\omega_i) = \frac{1}{m}\sum_{j=1}^{m} \hat{Z}_{i,j}$$

In the outer level of the simulation, we approximate the distribution of L by an empirical distribution. We assume $\omega_1, \ldots, \omega_n$ are n independent and identically distributed samples drawn from the real-world distribution of ω. Then we can approximate the different risk measurements by formula (2.3), (2.4) and (2.5).

2.3.1. MSE Estimator

In the uniform sampling, the total number of simulated samples is fixed as k, i.e. $k = nm$. The objective is to choose optimal allocation of $n \; and \; m$ so that we can minimize the mean squared error of estimators for different risk measurements. Gordy and Juneja (2006,2008) provided an analysis for the asymptotic behavior of the MSE of estimators for probability of large loss, Value-at-Risk and Expected Shortfall. The optimal uniform estimators for three measures are consistent. The optimal solution would utilize a number of outer scenarios n of order $k^{2/3}$, a number of inner samples m of order $k^{1/3}$, and result in an MSE of order $k^{-2/3}$.

2.4. Sequential Sampling

Broadie, Du and Moallemi (2010) have proposed a new algorithm called Efficient Sequential Sampling based on Gordy and Juneja (2008)'s work. It focused on one risk measure, the probability of large loss. The main idea of Sequential Sampling is sequentially allocating computational effort in the inner step simulation based on some marginal changes in the risk estimator for each outer scenario.

In uniform sampling, we uniformly distribute a constant number of inner samples for each outer scenario, which is straight-forward but may not be an efficient one. We consider now there are two outer scenarios ω_1 and ω_2 with the corresponding portfolio losses $L(\omega_1)$ and $L(\omega_2)$. Suppose that $L(\omega_1)$ is far away from the loss threshold u(much larger than u or less than u), and $L(\omega_2)$is quite near u. Denote $\hat{L}_1$ and $\hat{L}_2$ as the estimated portfolio loss under scenarios ω_1 and ω_2. We assume the estimated portfolio loss $\hat{L}_1$ and $\hat{L}_2$ follow the probability

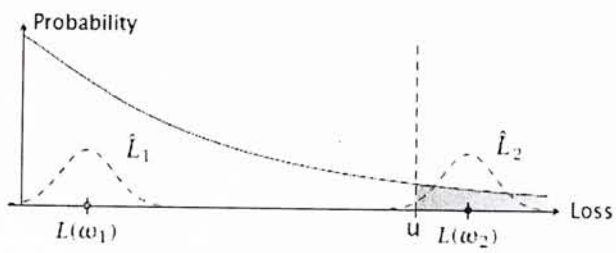

Figure 2.2: The motivation of non-uniform sampling.

distribution displayed in Figure 3.2. It is obvious that the $\hat{L}_1$ obtained by simulation is unlikely to impact the overall estimator $\hat{\alpha}$. However, in scenario ω_2, it is hard to determine whether $\hat{L}_2$ is larger or smaller than u. Thus we should allocate more inner steps in scenario ω_2, and fewer inner steps in scenario ω_1.

Broadie, Du and Moallemi (2010) developed an algorithm implementing the above idea. Then a modification method is followed to give asymptotic analysis for MSE of the estimator. After the discussion, Sequential Sampling can achieve a MSE of order $k^{-4/5}$. The optimal allocation is that outer scenario number n is of order $k^{4/5}$ and an average inner stage samples $\overline{m}$ is of order $k^{1/5}$.

CHAPTER 3

Methodology: Our Approach

3.1. Least-Squares Monte-Carlo Approach

In this work, we focus on risk measurement problem based on simulation. From previous discussions, we separate the simulation procedures into two parts. Firstly we simulate several outer scenarios to approximate the distribution of portfolio loss at time H. Secondly, we need to construct inner simulation process to approximate the portfolio value at time H as a conditional expectation in (2.2). In this way we can calculate the corresponding loss value for each outer scenario.

In pricing American option problem, it involves a comparison of the conditional expectation as the continuation value of the option and the exercise value of the option during the backward dynamic programming process. Longstaff and Schwartz (2001) proposed an influential algorithm, called Least Square Monte Carlo (LSM), in which the inner-level samples were used to estimate a parametric relationship between the state variables and the conditional expectation as the continuation value of the option. Following the similar idea of LSM, we may develop a Least Square approach in our risk measurement problem.

3.1.1. Framework

In the previous discussion, we have already known that to obtain the estimation of different risk measurements, we need to determine the distribution of our portfolio loss. However, it is generally hard to determine the exact value of the loss. Thus one important task is to evaluate the portfolio loss by monte-carlo simulation.

As we mentioned in Section 2.1, the formulas for the market value of the portfolio and portfolio loss at some future time H are

$$V_H = \sum_{k=1}^{K} V^{(k)}(H) = E^Q(\sum_{k=1}^{K} \int_H^{T_k} e^{-r(s-H)} dC^{(k)}(s)|\mathcal{F}_H); \tag{3.1}$$

$$L = V_0 - V_H = \sum_{k=1}^{K} V^{(k)}(0) - E^Q(\sum_{k=1}^{K} \int_H^{T_k} e^{-r(s-H)} dC^{(k)}(s)|\mathcal{F}_H).$$

Here V_0 is assumed to be known. To obtain the value of loss, we may approximate the conditional expectation by applying least-square regression on a suitable finite set of basis functions. In the paper "An analysis of a least squares regression method for American option pricing", Clément, Lamberton and Protter pointed out that the algorithm for LSM involved two steps of approximation. The first step of approximation was to replace the conditional expectation by a finite linear combination of "basis" functions. Within the second step of the approximation, they applied Monte-Carlo simulations and least squares regression to compute the linear combination given in approximation one. Our method also based on the two-step approximation methodology.

In the first approximation step We represent the conditional expectation V_H as a linear combination of a countable set of $\mathcal{F}_H$-measurable basis functions

$$V_H \approx \hat{V}_H^M(S_H) = \sum_{i=1}^{M} \beta_i \cdot f_i(S_H), \tag{3.2}$$

where the sequence $f_1(S_H), \ldots, f_M(S_H)$ are linearly independent in the Hibert space L^2. Thus it is reasonable to write an element of the L^2 space as a linear combination of the elements of the basis.

In the second approximation step We try to determine the coefficient value $\beta = (\beta_1, \ldots, \beta_M)$ using Monte-Carlo simulations and least-squares regression method.

The procedures are as follows.

- Firstly, we simulate n scenarios. Draw n simulation paths S_t for $t \in (0, H]$ under the physical measure. We denote these n scenarios by $\omega_1, \ldots, \omega_n$. Thus we obtain $S_H(\omega_1), \ldots, S_H(\omega_n)$.
- For each scenario ω_i, we simulate one path from time H to maturities under the risk-neutral measure. Whatever the number of cashflows is, we can always calculate the discounted cash flows along the simulated path. Then we can obtain the estimations for $V_H(\omega_1), \ldots, V_H(\omega_n)$ by formula (3.1).
- Subsequently, we use these observations to determine the coefficients β in the first approximation step by regression.

$$\hat{\beta}^{(n)} = \underset{\beta}{\operatorname{argmax}} \sum_{i=1}^{n} [V_H(\omega_i) - \sum_{j=1}^{M} \beta_j \cdot f_j(S_H(\omega_i))]^2$$

- Replacing β in formula (3.2) by $\hat{\beta}^{(n)}$, we achieve the second approximation

$$V_H \approx \hat{V}_H^M(S_H) \approx \hat{V}_H^{M,n}(S_H) = \sum_{i=1}^{M} \hat{\beta}_i^{(n)} \cdot f_i(S_H). \tag{3.3}$$

Thus for scenario ω_i, the LSM estimate for V_H is

$$V_H(\omega_i) \approx \hat{V}_H^{M,n}(S_H(\omega_i)) = \sum_{i=1}^{M} \hat{\beta}_i^{(n)} \cdot f_i(S_H(\omega_i)). \quad i = 1, \ldots, n \tag{3.4}$$

We can approximate the portfolio loss by $\hat{L}(\omega_i) = V_0 - \hat{V}_H^{M,n}(\omega_i)$, i=1,...,n. Based on these realizations, we can approximate the corresponding risk measurement.

3.2. Stochastic Mesh Method in risk measurement

By using two-level simulation to measure risk, Uniform Sampling and Sequential method both develop a tree shape in simulation. First we simulate n outer scenarios, then for each scenario, we simulate several inner paths. In American option pricing problem, to estimate the continuous value of the option, tree-building method is considered as the most straight-forward way. But if there are total k phases of periods until maturity, and for each phase, there are n number of scenarios to be generated, then the total paths will be n^k. To avoid the exponentially increase of total paths, Bordie and Glasserman proposed the stochastic mesh method which generated only n paths in total along the whole period. The same idea can be implemented in our risk measurement problem in the inner pricing level to estimate the portfolio value as a conditional expectation.

3.2.1. Framework

We start with a simplified situation: There is only one stream of cash flow at time T and $T > H$. The cash flow depends on the state variable S_T, and we write the payoff function as $V_T = h_T(S_T)$. We first give the simulation procedures.

First level for scenarios generating: Simulate n paths for S_t from time 0 to time H under physical measure. We denote these n scenarios by $\omega_1, \ldots, \omega_n$.

Second level for pricing the portfolio value: For each simulated scenario $S_H(\omega_i)$, we simulate one path from time H to maturities under the risk-neutral measure.

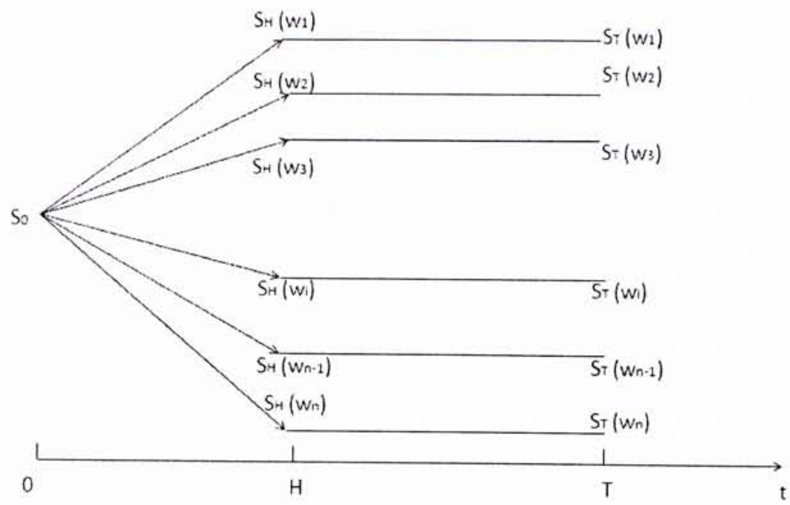

Figure 3.1: Simulation framework for S_t

Thus the portfolio value at time H is the discounted future cash flow

$$V_H(S_H(\omega)) = E^Q(h_T(S_T)e^{-r(T-H)}|S_H(\omega)). \tag{3.5}$$

Now the goal is to estimate the above conditional expectation by using all information at points $S_T(\omega_1), S_T(\omega_2), \ldots, S_T(\omega_n)$. The basic idea of stochastic mesh method is that we want to estimate the portfolio value at a node at time H by using values from all nodes at time T, not just the successors of current node. We may write the conditional expectation in (3.5) as a weighted average form.

$$\begin{aligned} V_H(S_H(\omega_i)) &= E^Q(h_T(S_T)e^{-r(T-H)}|S_H(\omega_i)) \\ &\approx \sum_{j=1}^{n} h_T(S_T(\omega_j))e^{-r(T-H)} \cdot w(T, S_H(\omega_i), S_T(\omega_j)). \end{aligned}$$

where $w(T, S_H(\omega_i), S_T(\omega_j))$ is a weight attached to the arc joining point $S_H(\omega_i)$ to point $S_T(\omega_j)$.

According to the above idea, we can estimate the portfolio value as an conditional expectation with n simulation paths in total. That is the main advantage and motivation behind stochastic mesh method. Then we will describe the detail implements of stochastic mesh method. We introduce a Markov chain $X_0, X_{t_1}, X_{t_2}, \ldots, X_T$ here.

- Generate the Mesh Points

  The first step is to generate the mesh points $X_t(i)$ for $i = 1, 2, \ldots, n$ and $t = H, T$.

  In the outer-level simulation, we already draw n paths for S_H under the physical measure. We set $X_H(i) = S_H(\omega_i)$ for $i = 1, 2, \ldots, n$, and generate $X_T(i)$ following some density function. The choice of density function will be described later.

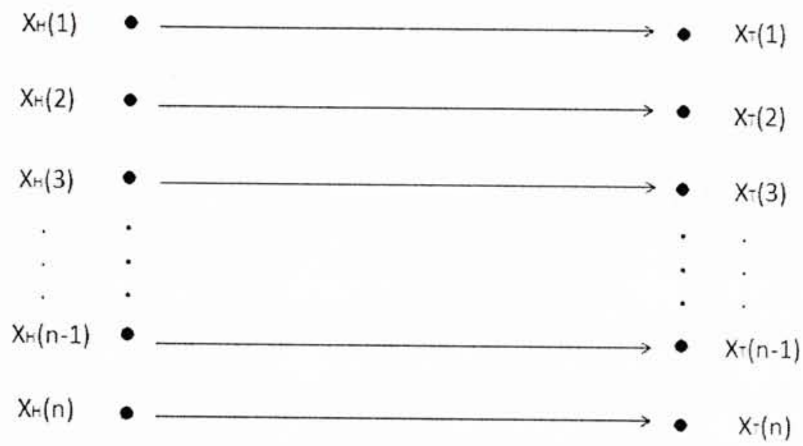

Figure 3.2: Generate the mesh points

Now we will estimate the portfolio value at time H as a weighted average of the mesh estimators at time T.

The mesh estimator is calculated by the following formulas

for $t = T$,

$$\hat{V}_T(X_T(i)) = h_T(X_T(i)),$$

where h_T is the payoff function at time T, and $i = 1, \ldots, n$;

for $t = H$,

$$\hat{V}_H(X_H(i)) = \frac{1}{n}\sum_{k=1}^{n} e^{-r(T-H)}\hat{V}_T(X_T(k)) \cdot w(T, X_H(i), X_T(k)), \tag{3.6}$$

where $w(T, X_H(i), X_T(k))$ is a weight attached to the arc joining mesh point $X_H(i)$ to mesh point $X_T(k)$.

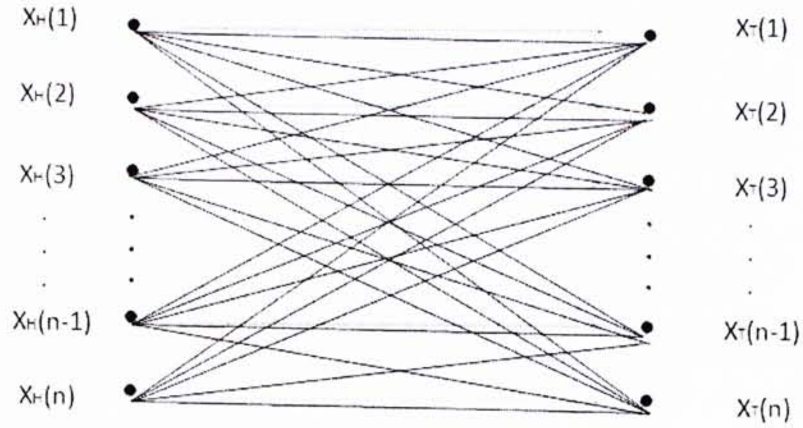

Figure 3.3: An illustration of calculation among mesh points

- Determine the Weights on the Arc.

  Suppose that conditional on $S_H = x$, we define the transition density $f_T(x, \cdot)$ for S_T as follows.

$$P(S_T \in A | S_H = x) = \int_A f_T(x, u) du,$$

  and $F_t(\cdot)$ denote the marginal density of S_t. We denote $g_t(\cdot)$ as the density function for X_t; $t = H, T$.

  Our approximation in (3.6) involves using $\hat{V}_T(X_T(k))$ for k=1,..., n, the information about the mesh points. But $X_T(k)$ for $k = 1, \dots, n$ are generated from the density function $g_T(\cdot)$ rather than $F_T(\cdot)$.

  Consider the portfolio value at time H

$$\begin{aligned}
V_H(S_H) &= E^Q(e^{-r(T-H)} V_T(S_T) | S_H = x) \qquad (3.7) \\
&= e^{-r(T-H)} \int V_T(u) \cdot f_T(x, u) du \\
&= e^{-r(T-H)} \int V_T(u) \cdot \frac{f_T(x, u)}{g_T(u)} g_T(u) du \\
&\equiv E^Q[e^{-r(T-H)} V_T(X_T(k)) \cdot \frac{f_T(x, X_T(k))}{g_T(X_T(k))}].
\end{aligned}$$

  From the last expression, we see that we can use information at the mesh points $X_T(k)$ to estimate $E(e^{-r(T-H)} V_T(S_T) | S_H = x)$ even though the points $X_T(k)$, $k = 1, 2, \dots, n$ are generated by the density $g_T(\cdot)$.

Comparing equations (3.6) and (3.7), we can make a conclusion that

$$\hat{V}_H(X_H(i)) = \frac{1}{n}\sum_{k=1}^{n}[e^{-r(T-H)}\ \hat{V}_T(X_T(k)) \cdot w(T, X_H(i), X_T(k))] \tag{3.8}$$

where

$$X_H(i) = S_H(\omega_i)\ \ for\ \ i = 1, 2, \ldots, n; \quad \hat{V}_T(X_T(k)) = h_T(X_T(k))\ \ for\ \ k = 1, 2, \ldots, n$$

and

$$w(T, X_H(i), X_T(k)) = \frac{f_T(X_H(i), X_T(k))}{g_T(X_T(k))}.$$

- Determine the Density Function for Mesh Points

  Up to now, we haven't discuss the probability density function by which the random mesh points are generated. In Broadie and Glasserman (2004), to avoid exponential growth in variance of the estimator, they proposed one way to choose the density $g_t(\cdot)$ as the average transition density

$$g_t(u) = \frac{1}{n}\sum_{j=1}^{n} f_t(X_{t-1}(j), u)$$

  i.e.

$$g_T(u) = \frac{1}{n}\sum_{j=1}^{n} f_T(X_H(j), u)$$

  where $f_T(x, \cdot)$ is the transition density for S_T given that $S_H = x$.

  Thus the weight function $w(T, X_H(i), X_T(k))$ may be written as

$$w(T, X_H(i), X_T(k)) = \frac{f_T(X_H(i), X_T(k))}{\frac{1}{n}\sum_{j=1}^{n} f_T(X_H(j), X_T(k))}.$$

  In our setting, no intermediate cashflows is involved, and only one step exists from H to T, there is no exponential growth variance problem. Thus another nature choice for $g_t(\cdot)$ is to set $g_t(u) = f_t(u)$, the marginal density for S_t.

3.2.2. With a series of cash flows

We have discussed the situation where there is only one cash flow between time interval $(H, T]$. Now we consider the case that there are cashflows at time points $(t_1, t_2, \ldots, t_i, \ldots, t_N = T)$ and $H = t_0 < t_1 < \cdots < t_i < t_N = T$. We assume there exists a series of payoff functions $(h_{t_1}(S_{t_1}), h_{t_2}(S_{t_2}), \ldots, h_{t_N}(S_{t_N}))$.

We still build up the mesh points $X_t(j)$ for $j = 1, 2, ,, n$ and $t = t_0, \ldots, t_N$. Set $X_{t_0}(j) = S_H(\omega_j)$ for $j = 1, 2, \ldots, n$ and generate only one path $X_{t_i}(j)$ from $X_{t_0}(j)$ using some density function $g_{t_i}(\cdot)$.

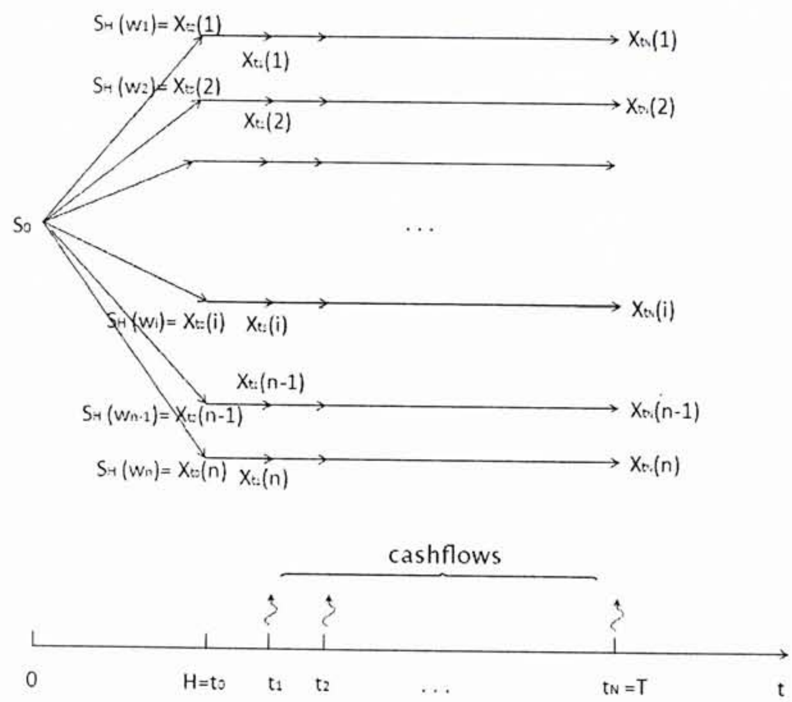

Figure 3.4: An illustration of multiple cash flows in the future

Thus the portfolio value at time H equals to the cumulative value of future discounted cashflows.

$$
\begin{aligned}
V_H(S_H = x) &= E^Q[\int_H^T e^{-r(s-H)} dC(s)|S_H = x] \\
&= E^Q[\sum_{i=1}^{N} e^{-r(t_i-t_0)} h_{t_i}(S_{t_i})|S_{t_0} = x] \\
&= \sum_{i=1}^{N} E^Q[e^{-r(t_i-t_0)} h_{t_i}(S_{t_i})|S_{t_0} = x].
\end{aligned}
$$

We can treat V_H as a summation of a series of conditional expectations. Subsequently, we apply the above simplified stochastic mesh method for each term. In more detail,

for $t = t_i$

$$
E^Q[e^{-r(t_i-t_0)} h_{t_i}(S_{t_i})|S_{t_0}(\omega_k)] \approx \frac{1}{n}\sum_{j=1}^{n} e^{-r(t_i-t_0)}\ h_{t_i}(X_{t_i}(j)) \cdot w(t_i, X_{t_0}(k), X_{t_i}(j)).
$$

where

$$
w(t_i, X_{t_0}(k), X_{t_i}(j)) = \frac{f_{t_i}(X_{t_0}(k), X_{t_i}(j))}{g_{t_i}(X_{t_i}(j))}.
$$

Thus

$$
\begin{aligned}
\hat{V}_H(S_H(\omega_k)) &= \hat{V}_{t_0}(X_{t_0}(k)) \qquad (3.9)\\
&= \frac{1}{n}\sum_{i=1}^{N}\sum_{j=1}^{n} e^{-r(t_i-t_0)}\ h_{t_i}(X_{t_i}(j)) \cdot w(t_i, X_{t_0}(k), X_{t_i}(j)).
\end{aligned}
$$

3.2.3. Derive Marginal Density and Transition Density

In this section, we derive marginal density and transition density for stochastic mesh method. We first specify the following basic settings: Recall that S_t is the D-dimensional vector of several state variables that govern portfolio prices. Denote $S_t = (S_t^1, S_t^2, \ldots, S_t^D)$. Without loss of generality, we assume the underlying S_t is a D-dimensional, m-factor lognormal processes with the real-world drift μ_d:

$$
\frac{dS_t^d}{S_t^d} = \mu_d dt + \sum_{j=1}^{m} L_{dj} dW_t^j, \quad d = 1, \ldots, D,
$$

where W_t^j for $j = 1, \dots, m$ are independent standard Brownian motions; L is an $D \times m$ matrix. The instantaneous covariance matrix is $\Sigma = LL^\top$.

Thus

$$S_t^d = S_0^d \exp[(\mu_d - \frac{1}{2}\Sigma_{dd})t + \sqrt{t}\sum_{j=1}^{m} L_{dj} W_t^j] \qquad for \;\; d = 1, 2, \dots, D.$$

Theorem 3.1. *Suppose S_t follows Geometric Brownian motion with μ as the real-world drift and Σ as the covariance matrix. r is the risk-neutral rate. We simulate S_t from time 0 to time H under physical measure and simulate S_H to S_T from risk-neutral measure. Then we have*

$$\ln S_T \sim N\big(\bar{S}_1, \Sigma T\big),$$

where $\bar{S}_1 \triangleq \ln S_0 + (\mu - \frac{1}{2}diagonal(\Sigma))H + (r - \frac{1}{2}diagonal(\Sigma))(T-H)$. Hence the density function of $\ln S_T$ is

$$f_1(u) = \frac{1}{(2\pi)^{D/2} det(\Sigma T)^{1/2}} \exp\{-\frac{1}{2}(u - \bar{S}_1)^T (\Sigma T)^{-1} (u - \bar{S}_1)\}.$$

The density function of S_T is $F(u) = f_1(\ln u)/u_1 u_2 \dots u_D$ where u is a D-dimensional vector and u_i is the ith element of vector u.

The proof of Theorem 3.1 is provided in Appendix A.

Theorem 3.2. *Under the above setting, given $S_H = x$, we simulate S_T from S_H under risk-neutral measure. Thus, the D-dimensional vector $\ln S_T$ follows multivariate normal distribution.*

$$\ln S_T \sim N\big(\bar{S}_2, \Sigma(T-H)\big),$$

where $\bar{S}_2 \triangleq \ln x + (r - \frac{1}{2}diagonal(\Sigma))(T-H)$.

Given $\ln S_H = \ln x$, the density function for $\ln S_T$ is

$$f_2(x, u) = \frac{1}{(2\pi)^{D/2} det(\Sigma(T-H))^{1/2}} \exp\{-\frac{1}{2}(u - \bar{S}_2)^T (\Sigma(T-H))^{-1} (u - \bar{S}_2)\}.$$

Thus the transition density of S_T given $S_H = x$ is $f_T(x, u) = f_2(x, \ln u)/u_1 u_2 \dots u_D$ where u is a D-dimensional vector and u_i is the ith element of vector u.

The proof is provided in Appendix A.

CHAPTER 4

Numerical Experiments

In this chapter we present several numerical experiments that compare the four methods to compute risk measurements under one-dimensional and multi-dimensional situations. We begin in section 4.1 by describing our experimental settings. In section 4.2 and 4.3, we compare the bias and MSE of the estimators obtained by each approach. Finally, we make some conclusions on the numerical results.

4.1. Experimental Setting

In our experiment setting, $S_t = (S_t^1, S_t^2, \ldots, S_t^D)$ is a D-dimensional vector of several state variables that governs portfolio prices. We assume S_t follow Geometric Brownian motion with μ as the real-world drift and Σ as the covariance matrix. For each element of S_t

$$\frac{dS_t^d}{S_t^d} = \mu_d dt + \sum_{j=1}^{m} L_{dj} dW_t^j, \quad d = 1, \ldots, D,$$

where W_t^j for $j = 1, \ldots, m$ are independent standard Brownian motions; L is an $D \times m$ matrix. The covariance matrix $\Sigma = LL^\top$.

Thus

$$S_t^d = S_0^d \, \exp[(\mu_d - \frac{1}{2}\Sigma_{dd})t + \sqrt{t}\sum_{j=1}^{m} L_{dj} W_t^j] \qquad for \ \ d = 1, 2, \ldots, D.$$

- One-Dimensional Case

  We assume the portfolio consists of a long position in a single put option with strike price $K = 95$ and maturity $T = 0.25$ years. Let the initial price of the underlying asset be $S_0 = 100$. The real drift of this process is $\mu = 8\%$ and volatility is $\sigma = 20\%$. The risk-free rate is $r = 3\%$. We fix the risk horizon H at one week, i.e. $H = 1/52$ years.

- Multi-Dimensional Case

  In the multi-dimensional case, we assume that S_t is a two dimensional lognormal process with initial prices $S_0 = [93\ 77]$; The real drifts $\mu = [0.08\ 0.03]$ and covariance matrix $\Sigma = \begin{pmatrix} 0.04 & 0.005 \\ 0.005 & 0.0025 \end{pmatrix}$. We consider a portfolio consisting of long two call options with strike prices $K = [90\ 75]$; maturities $T = [1.5\ 1.5]$. Different option is based on different underlying asset respectively. Let the risk horizon be four months, $H = 1/3$. And the risk-free rate is $r = 4\%$.

For these four methods, given a total computational budget k, we consider the following estimators.

- Uniform sampling. The optimal estimator with minimium MSE is $m \propto k^{1/3}$, $n \propto k^{2/3}$. In practice, we just simply choose $m = k^{1/3}$ and $n = k^{2/3}$.

- Sequential sampling. Given total computational budget k, we use initial inner samples as $m_0 = 10$, and average inner samples as $\overline{m} = 100$. Thus outer scenarios n is fixed by $n = k/\overline{m}$. It should be noticed the sequential method can only apply in estimating the probability of large loss due to the special characteristics of the derivation.

- Stochastic mesh method. In this method, we simulate n outer scenarios and for each outer scenario, we just simulate one inner sample. But we can not simply set $n = k$. Recall the mesh estimator calculation formula (3.5), for n outer scenarios, the total computational budget is approximately n^2. Thus we set $n = k^{1/2}$.

- Least Square method. The estimator of least square method depends on the selection of basis function. In our experiments, we use $\{1, S, S^2\}$ as basis functions in one-dimensional case. For multi-dimensional case, we choose basis functions consisting of a constant, each element of the state vector, their quadratic value and their cross products.

To keep consistency in time, we directly plot bias and MSE versus running time by varying the total budget k from 10^4 to 10^7.

4.2. Bias Comparison

- One-Dimensional case

  We first numerically compare the estimators obtained by the four methods on the basis of bias in one-dimensional case. Figure 4.1 shows us the bias comparison of estimators for estimating the probability of large loss. The results for Value-at-Risk and Expected Shortfall are presented in Figure 4.2 and Figure 4.3.

  In one-dimensional case, we can see that Stochastic Mesh method can achieve the smallest bias compared to other methods. For Least Square method, the bias decreases quickly at the beginning. However, as time goes the bias of Least Square method turns out to remain at a certain level. To conclude, the Least Square method converges quickly but it may not produce an accurate enough estimator. Since the choices of basis functions is crucial to the performance of Least Square method, we may add more basis functions to achieve smaller biases.

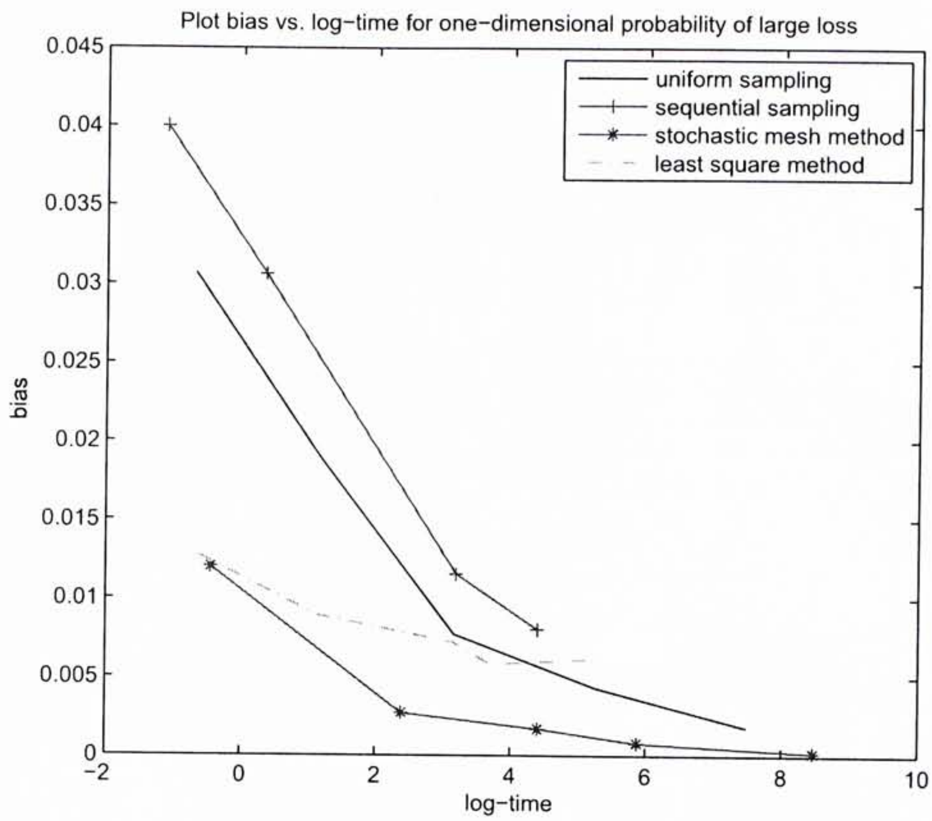

Figure 4.1: Bias plot for probability of large loss in 1-D case. Basic parameter for numerical illustration: Long position in a single put option with strike price $K = 95$; maturity $T = 0.25$ years; initial asset price $S_0 = 100$. The real drift of this process is $\mu = 8\%$, volatility is $\sigma = 20\%$ and risk-free rate is $r = 3\%$. We fix the risk horizon one week, $H = 1/52$ years. Stochastic Mesh method can achieve the smallest bias. For Least Square method, small error can be achieved within a short time. But the convergent error is still large.

- Multi-Dimensional case

  The bias performance for each method under multi-dimensional case is basically similar. Figure 4.4 shows us the bias plot for estimators of probability of loss in a multi-dimensional case. Then the bias comparison on the basis of Value-at-Risk is shown in Figure 4.5. At last, Figure 4.6 gives the bias plot for estimating Expected Shortfall. In this case, Least Square method and Stochastic Mesh both can achieve relative small errors.

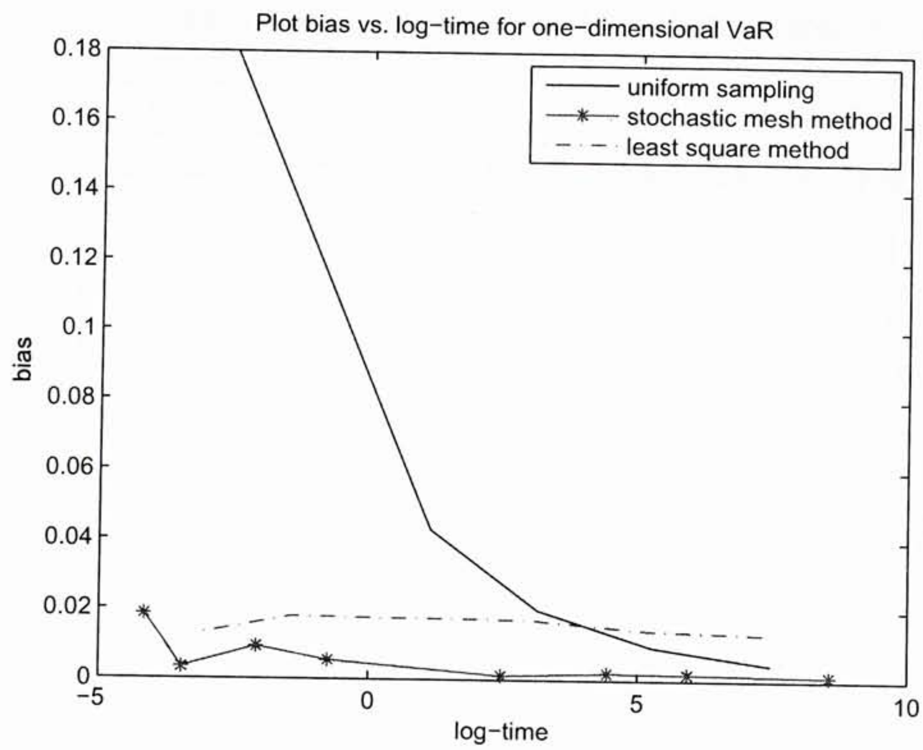

Figure 4.2: Bias plot for VaR in 1-D case. Basic parameter for numerical illustration: Long position in a single put option with strike price $K = 95$; maturity $T = 0.25$ years; initial asset price $S_0 = 100$. The real drift of this process is $\mu = 8\%$, volatility is $\sigma = 20\%$ and risk-free rate is $r = 3\%$. We fix the risk horizon one week, $H = 1/52$ years. Stochastic Mesh method can achieve the smallest bias. For Least Square method, small error can be achieved within a short time. But the convergent error is still large.

4.3. MSE Comparison

- One-Dimensional case

  In this section, we will provide numerically MSE comparison of various estimators obtained by the past two methods and our new approaches. At first, we conduct MSE comparison in a one-dimensional case. Figure 4.7, Figure 4.8 and Figure 4.9 present the plot of MSE versus time for probability of large loss, VaR and Expected Shortfall respectively.

  We can find out that, in MSE comparison, Stochastic Mesh method perform well under one-dimensional case. Least Square method can often attain a

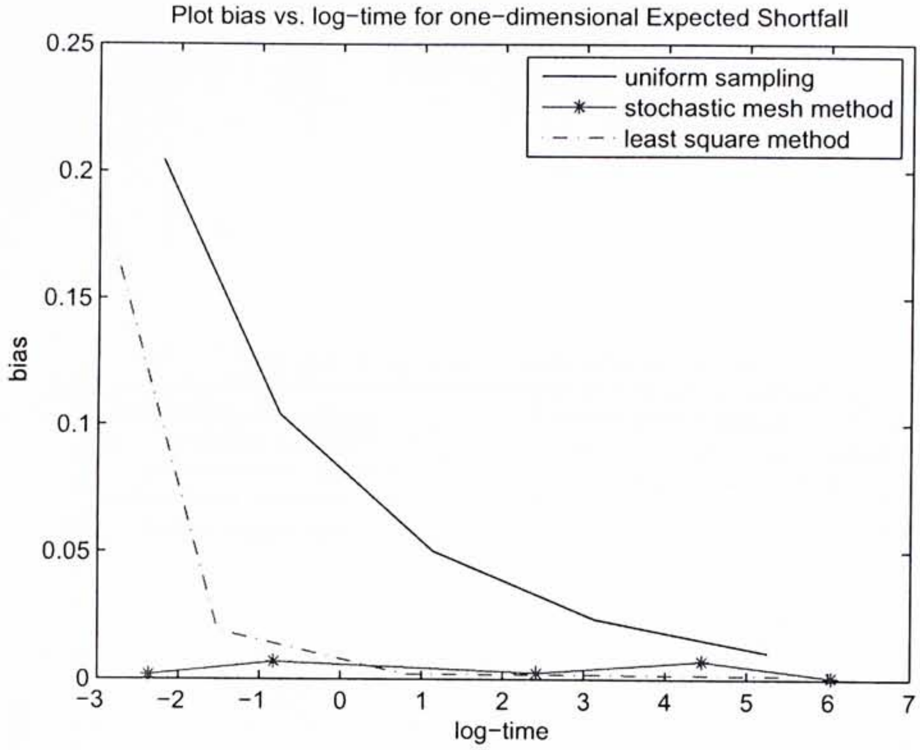

Figure 4.3: Bias plot for ES in 1-D case. Basic parameter for numerical illustration: Long position in a single put option with strike price $K = 95$; maturity $T = 0.25$ years; initial asset price $S_0 = 100$. The real drift of this process is $\mu = 8\%$, volatility is $\sigma = 20\%$ and risk-free rate is $r = 3\%$. We fix the risk horizon one week, $H = 1/52$ years. Stochastic Mesh method and Least Square method both can achieve the small bias.

smaller MSE in a few seconds. However, as time goes by, the MSE of Least Square method have relatively flat trend because it is dominated by the corresponding large bias term. Thus we can conclude that Least Square method converges fast. It can reach a small bias and MSE in a few seconds. On the other hand, the final bias level will be relatively large. Thus we may improve the bias performance of Least Square method by choosing proper basis functions.

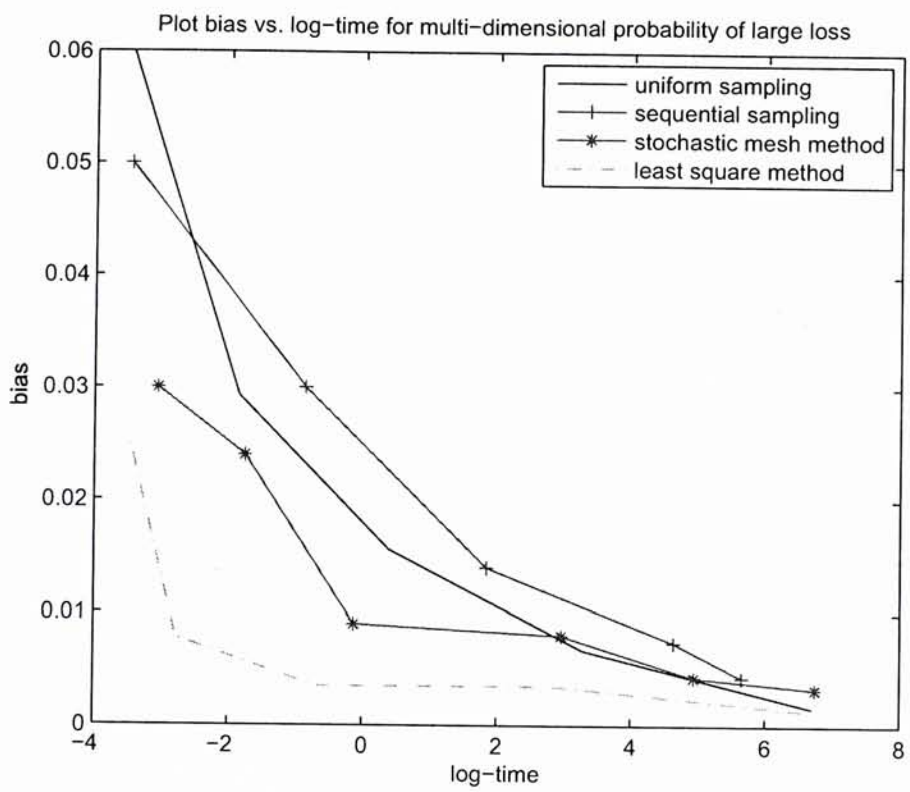

Figure 4.4: Bias plot for probability of large loss in multi-dimensional case. Basic parameter for numerical illustration: There are two assets with initial prices $S_0 = [93\ 77]$; The real drifts $\mu = [0.08\ 0.03]$ and volatilities are $\sigma_1 = 0.2$, $\sigma_2 = 0.05$ and the correlation coefficient $\rho = 0.5$. We consider a portfolio consisting of long three call options with strike prices $K = [90\ 75]$; maturities $T = [1.5\ 1.5]$. Risk horizon is $H = 1/3$; risk-free rate is $r = 4\%$.

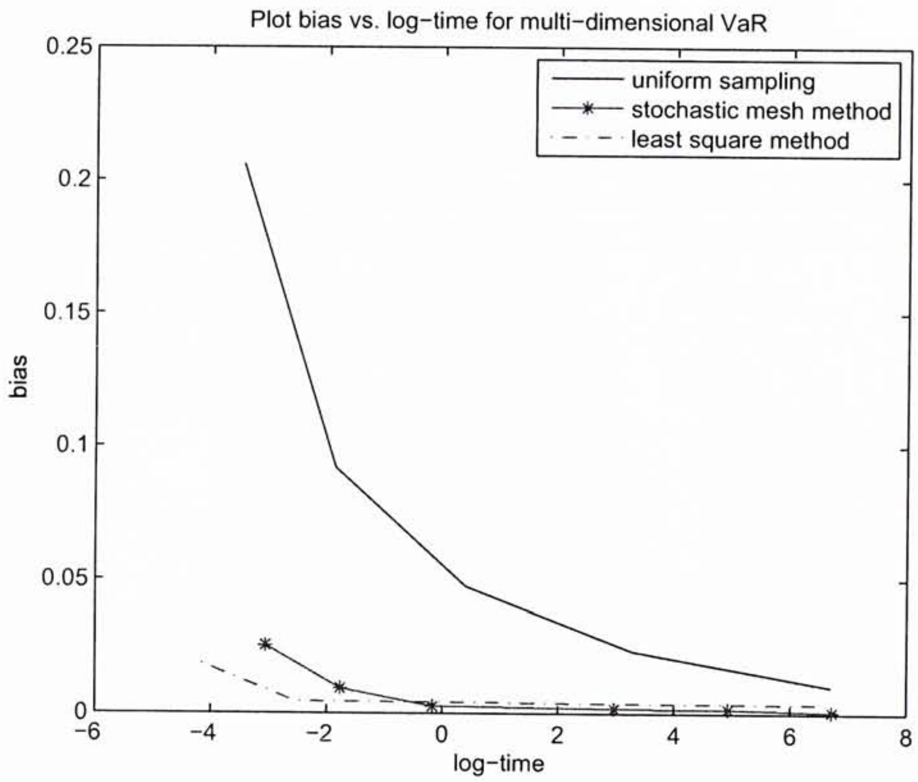

Figure 4.5: Bias plot for VaR in multi-dimensional case. Basic parameter for numerical illustration: There are two assets with initial prices $S_0 = [93\ 77]$; The real drifts $\mu = [0.08\ 0.03]$ and volatilities are $\sigma_1 = 0.2$, $\sigma_2 = 0.05$ and the correlation coefficient $\rho = 0.5$. We consider a portfolio consisting of long three call options with strike prices $K = [90\ 75]$; maturities $T = [1.5\ 1.5]$. Risk horizon is $H = 1/3$; risk-free rate is $r = 4\%$. We can see Least Square method and Stochastic Mesh both can achieve relative small errors.

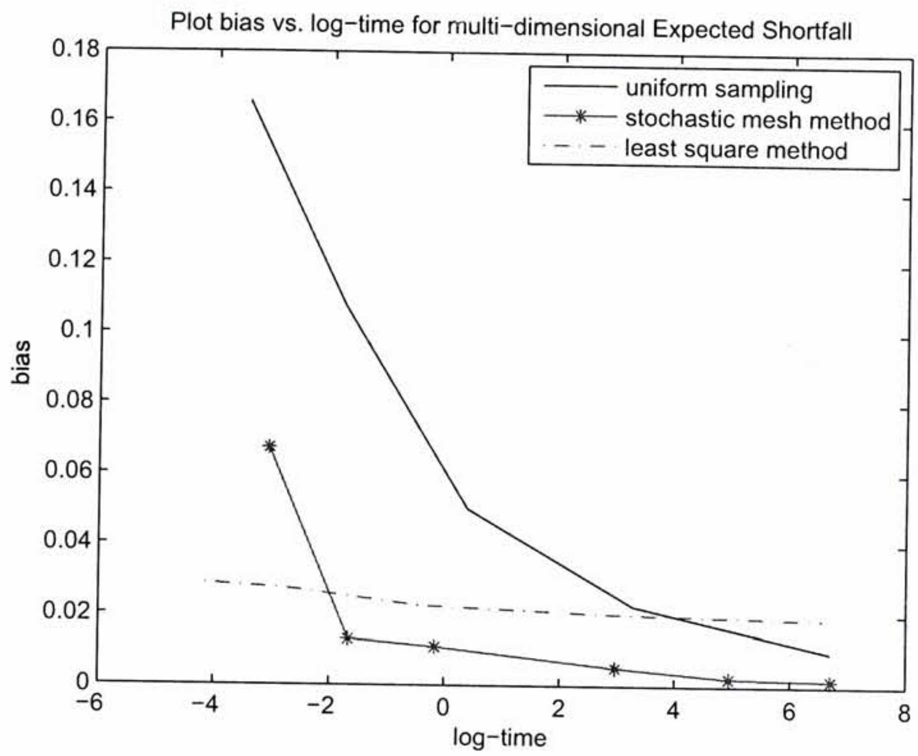

Figure 4.6: Bias plot for ES in multi-dimensional case. Basic parameter for numerical illustration: There are two assets with initial prices $S_0 = [93\ 77]$; The real drifts $\mu = [0.08\ 0.03]$ and volatilities are $\sigma_1 = 0.2$, $\sigma_2 = 0.05$ and the correlation coefficient $\rho = 0.5$. We consider a portfolio consisting of long three call options with strike prices $K = [90\ 75]$; maturities $T = [1.5\ 1.5]$. Risk horizon is $H = 1/3$; risk-free rate is $r = 4\%$. For Least Square method, small error can be achieved within a short time. Stochastic Mesh both can achieve smallest error among four methods.

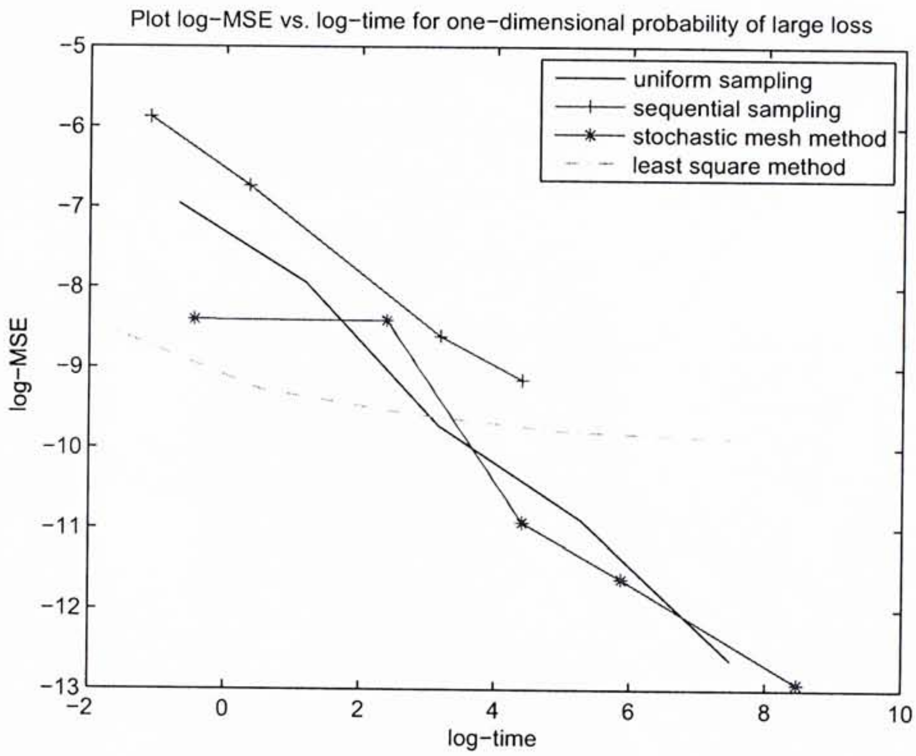

Figure 4.7: MSE plot for probability of large loss in 1-D case. Basic parameter for numerical illustration: Long position in a single put option with strike price $K = 95$; maturity $T = 0.25$ years; initial asset price $S_0 = 100$. The real drift of this process is $\mu = 8\%$, volatility is $\sigma = 20\%$ and risk-free rate is $r = 3\%$. We fix the risk horizon one week, $H = 1/52$ years.

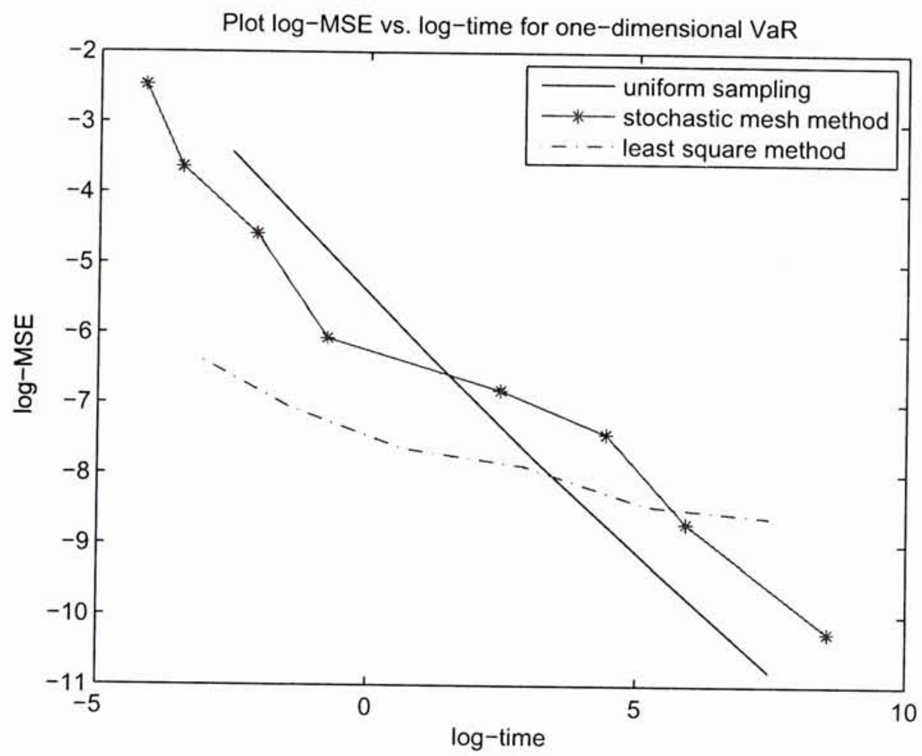

Figure 4.8: MSE plot for VaR in 1-D case. Basic parameter for numerical illustration: Long position in a single put option with strike price $K = 95$; maturity $T = 0.25$ years; initial asset price $S_0 = 100$. The real drift of this process is $\mu = 8\%$, volatility is $\sigma = 20\%$ and risk-free rate is $r = 3\%$. We fix the risk horizon one week, $H = 1/52$ years.

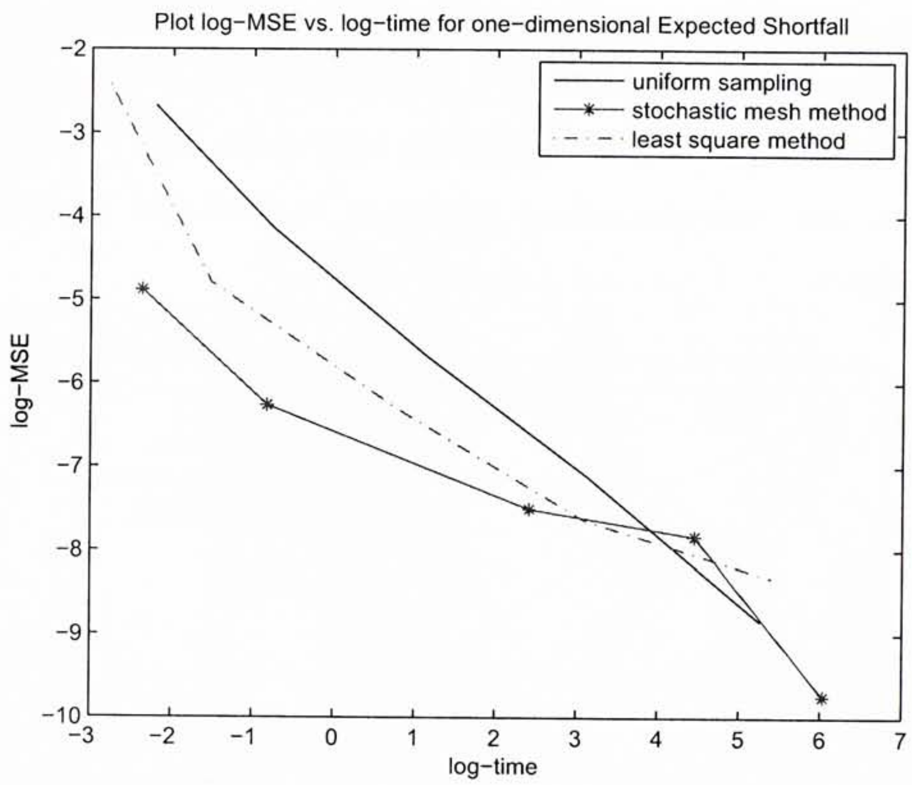

Figure 4.9: MSE plot for ES in 1-D case. Basic parameter for numerical illustration: Long position in a single put option with strike price $K = 95$; maturity $T = 0.25$ years; initial asset price $S_0 = 100$. The real drift of this process is $\mu = 8\%$, volatility is $\sigma = 20\%$ and risk-free rate is $r = 3\%$. We fix the risk horizon one week, $H = 1/52$ years.

- Multi-Dimensional case

The numerical results of MSE plot under multi-dimensional case are presented in Figure 4.10, Figure 4.11 and Figure 4.12. In this case, Least Square method converges quickly and incurs the smallest MSE in multi-dimensional cases. Unfortunately, the MSE performance of Stochastic Mesh method is not quite satisfactory. The reason may be that in multi-dimensional case, MSE of Stochastic Mesh method is dominated by the variance term. The multi-dimensional underlying assets may increase the computational complexity in calculating mesh estimators. Thus within a fixed time, less outer scenarios is allowed to be simulated, which incurs a large variance.

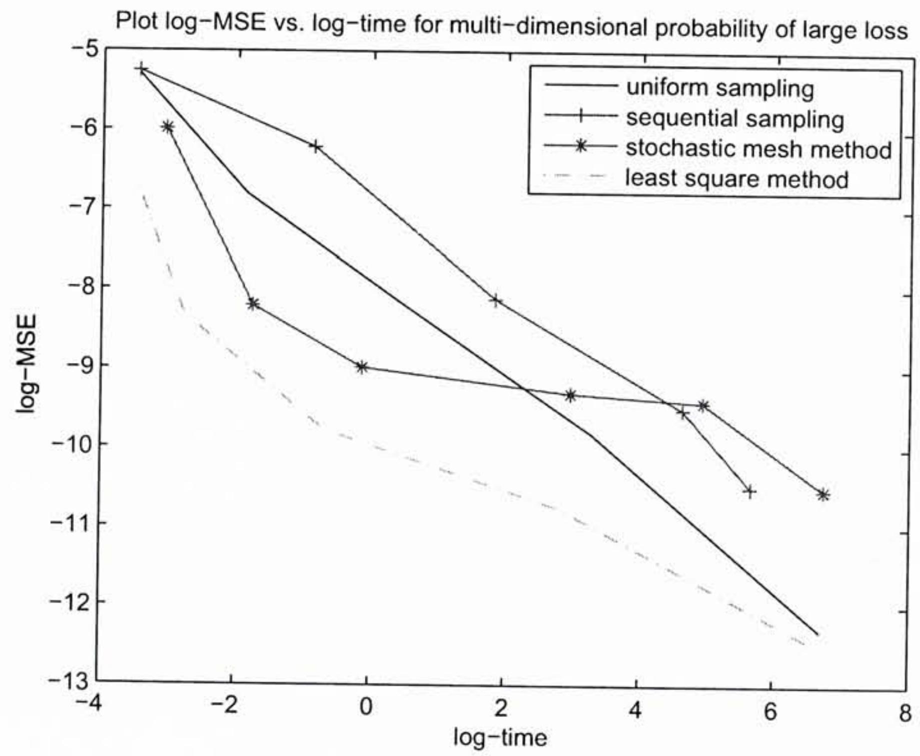

Figure 4.10: MSE plot for probability of large loss in multi-dimensional case. Basic parameter for numerical illustration: There are two assets with initial prices $S_0 = [93\ 77]$; The real drifts $\mu = [0.08\ 0.03]$ and volatilities are $\sigma_1 = 0.2$, $\sigma_2 = 0.05$ and the correlation coefficient $\rho = 0.5$. We consider a portfolio consisting of long three call options with strike prices $K = [90\ 75]$; maturities $T = [1.5\ 1.5]$. Risk horizon is $H = 1/3$; risk-free rate is $r = 4\%$.

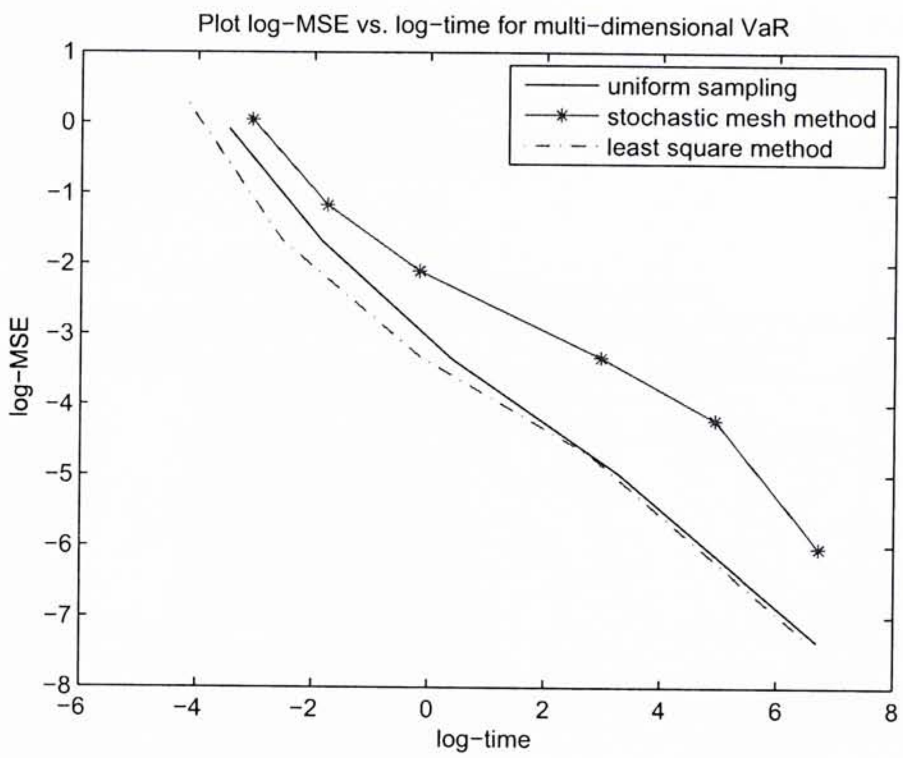

Figure 4.11: MSE plot for VaR in multi-dimensional case. Basic parameter for numerical illustration: There are two assets with initial prices $S_0 = [93\ 77]$; The real drifts $\mu = [0.08\ 0.03]$ and volatilities are $\sigma_1 = 0.2$, $\sigma_2 = 0.05$ and the correlation coefficient $\rho = 0.5$. We consider a portfolio consisting of long three call options with strike prices $K = [90\ 75]$; maturities $T = [1.5\ 1.5]$. Risk horizon is $H = 1/3$; risk-free rate is $r = 4\%$.

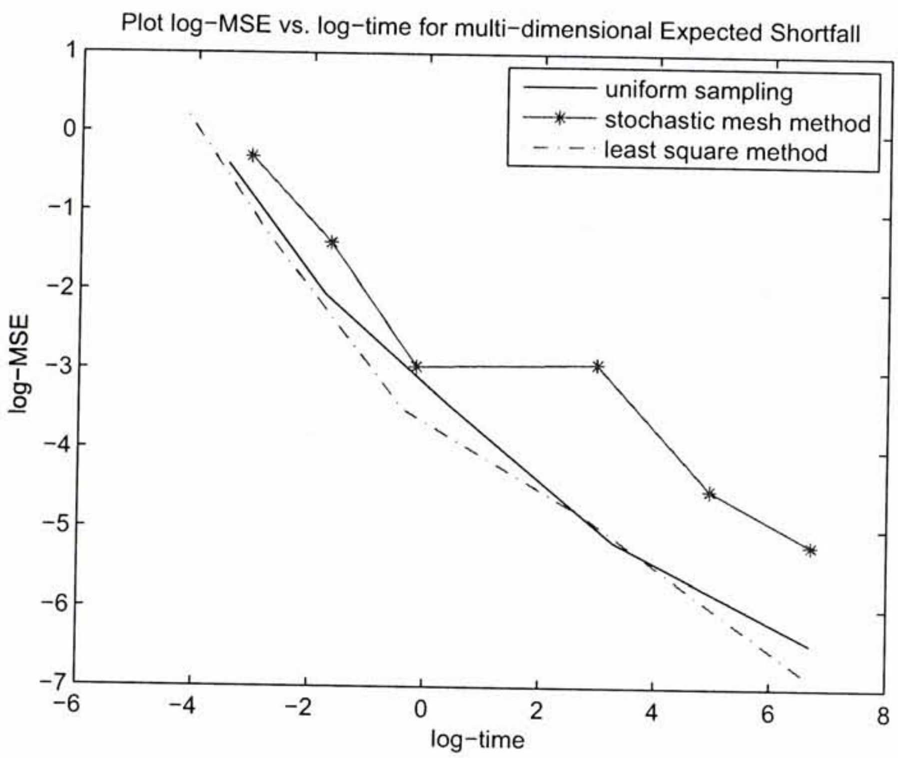

Figure 4.12: MSE plot for ES in multi-dimensional case. Basic parameter for numerical illustration: There are two assets with initial prices $S_0 = [93\ 77]$; The real drifts $\mu = [0.08\ 0.03]$ and volatilities are $\sigma_1 = 0.2$, $\sigma_2 = 0.05$ and the correlation coefficient $\rho = 0.5$. We consider a portfolio consisting of long three call options with strike prices $K = [90\ 75]$; maturities $T = [1.5\ 1.5]$. Risk horizon is $H = 1/3$; risk-free rate is $r = 4\%$.

4.4. Modified Least Square method

In the numerical experiments, we can see that Least Square method can achieve a smaller error within a short time period. However the converged bias seems to be relatively large. One solution to enhancing the bias performance may be adding more basis functions in our experiments. We retry the experiment of estimating Value-at-Risk in one-dimensional case. Now we choose 4 basis functions $\{1, S, S^2, S^3\}$ instead of $\{1, S, S^2\}$. The following figures shows the bias performance and MSE performance of modified Least Square method comparing with others.

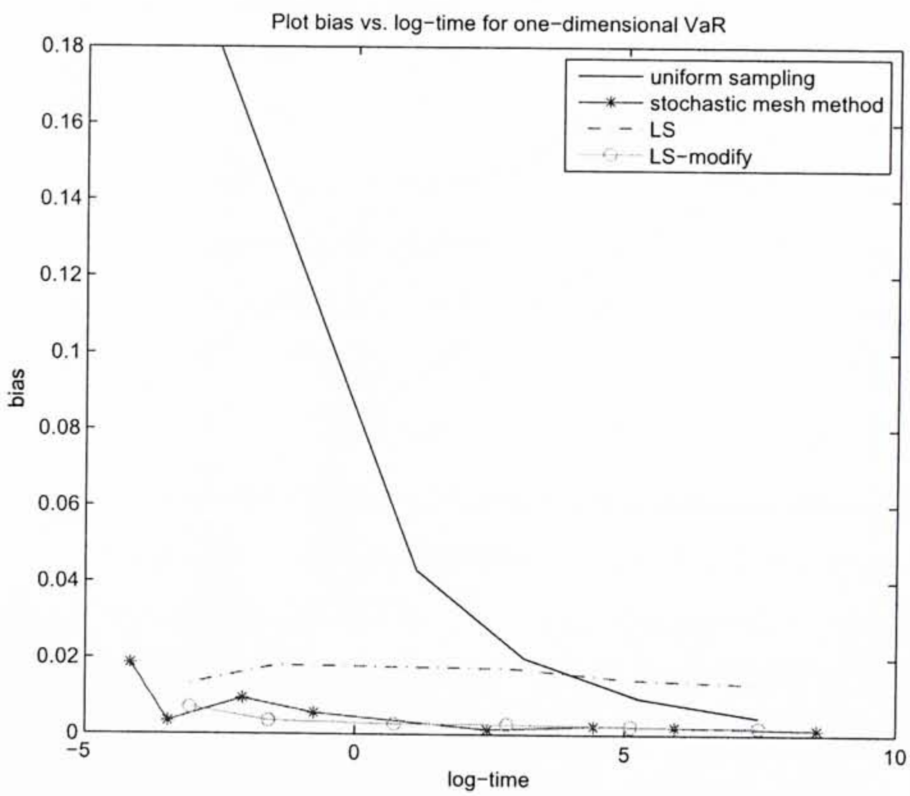

Figure 4.13: Bias performance of modified Least Square method for VaR in 1-D case. Modified Least Square method can achieve smaller bias than previous Least Square method.

From Figure (4.13) and Figure (4.14), we can see that the bias of modified Least Square method is better that the old one. The associated MSE also converges to an even smaller scale. To conclude, the choice of basis functions is

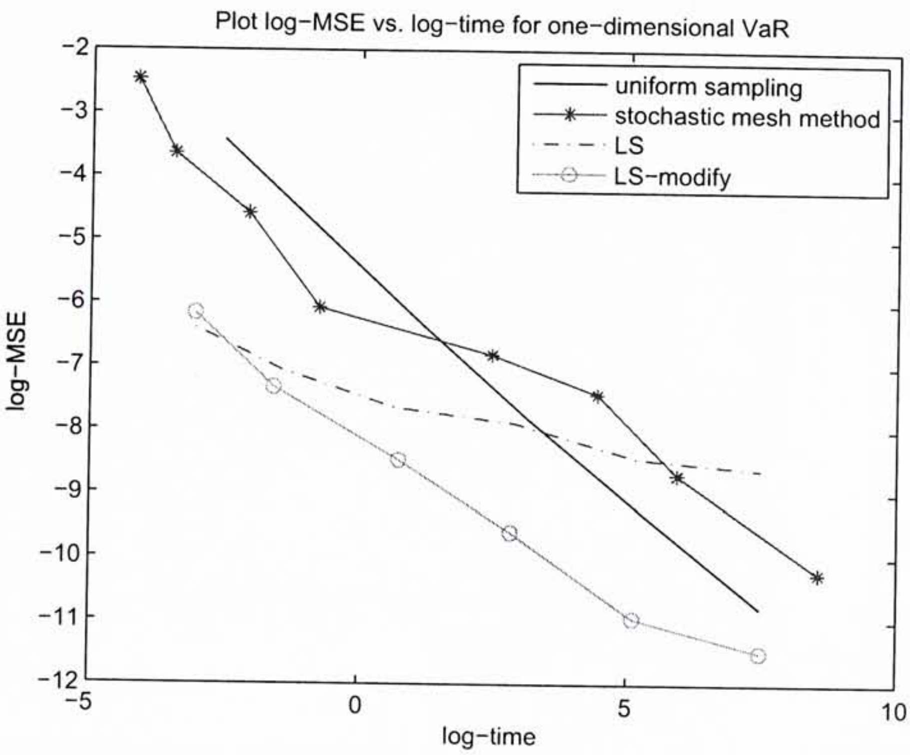

Figure 4.14: MSE performance of modified Least Square method for VaR in 1-D case. MSE performance of modified Least Square method has been improved.

crucial in Least Square method. By adding more basis functions, we can obtain more accurate estimators by Least Square method.

CHAPTER 5

CONCLUSION

In this thesis, we consider the problem of estimating three different risk measurements - Probability of Large Loss; Value-at-Risk; and Expected shortfall for portfolio by nested simulation procedures. Gordy and Juneja (2006, 2008) and Broadie, Du, and Moallemi (2010)'s methods have been reviewed first. Then, we develop two approaches - Least Square Monte-Carlo approach and Stochastic Mesh approach under our setting, which used to apply to pricing American style option. We use these two methods to estimate the future portfolio value as a conditional expectation in the inner level simulation step. In the numerical experiment part, we conduct several numerical examples and compare the four approaches numerically in bias comparison and MSE comparison. We find out that Uniform sampling method mentioned by Gordy and Juneja has its limitation in practice. It is not clear how to determine the coefficient for optimal MSE estimators. Meanwhile, due to the special characteristics of the derivation, Sequential Sampling proposed by Broadie, Du, and Moallemi can not be applied to estimate other risk measurements except for the probability of large loss. From the numerical experiments, we can see that our two new approaches have advantages in bias comparison in both one-dimensional case and multi-dimensional case. Least Square method can be fast and easy to implement without allocating computational budgets into outer and inner level simulations. There are some concerns about the bias of the least square method. The converging error may remain at a relatively higher level. However, by choosing proper basis function for the least square method, the bias can be effectively reduced. Stochastic Mesh method incurs the smallest error but is time-consuming in multi-dimensional cases. In comparing MSE experiments, Stochastic Mesh method can not perform as well as uniform sampling and Least Square method in multi-dimensional case. The reason may be that in multi-dimensional case, MSE of Stochastic Mesh method is dominated by the variance term. The multi-dimensional underlying assets may increase the computational complexity in calculating mesh estimators. Thus within a fixed time, less outer scenarios is allowed to be simulated, which incurs a large variance.

Since we focus on the numerical implements of each method with little theoretical work, in the future, we may approach to give some analytical results about the convergence rate of MSE under our two new methods. Also, to improve the efficiency of Least Square method, we may focus more on the selection of basis function in Least-Square method. Moreover, a weighted least square regression approach may also be considered.

APPENDIX A

Appendix A

A.1. Proof of Theorem 3.1

S_t is a D-dimensional, m-factor lognormal processes with the real-world drift μ_d:

$$\frac{dS_t^d}{S_t^d} = \mu_d dt + \sum_{j=1}^{m} L_{dj} dW_t^j, \quad d = 1, \ldots, D,$$

where W_t^j for $j = 1, \ldots, m$ are independent standard Brownian motions; L is an $D \times m$ matrix. The instantaneous covariance matrix is $\Sigma = LL^\top$.

Thus

$$S_t^d = S_0^d \, \exp[(\mu_d - \frac{1}{2}\Sigma_{dd})t + \sqrt{t}\sum_{j=1}^{m} L_{dj} W_t^j] \qquad for \;\; d = 1, 2, \ldots, D.$$

$$\ln S_T^d = \ln(\frac{S_T^d}{S_H^d}) + \ln S_H^d$$

where $\ln(\frac{S_T^d}{S_H^d})$ and $\ln S_H^d$ are both normal and independent from each other.

And

$$S_H^d = S_0^d \exp\{(\mu_d - \frac{1}{2}\Sigma_{dd})H + \sum_{j=1}^{m} L_{dj} W_H^j\}, \qquad d = 1, \ldots, D \tag{A.1}$$

$$\ln S_H^d = \ln S_0^d + (\mu_d - \frac{1}{2}\Sigma_{dd})H + \sum_{j=1}^{m} L_{dj} W_H^j, \qquad d = 1, \ldots, D$$

$$\ln S_H^d \sim N\big(\ln S_0^d + (\mu_d - \frac{1}{2}\Sigma_{dd})H, \Sigma_{dd}H\big), \quad d = 1, \ldots, D$$

The D-dimensional vector $\ln S_H$ follows multivariate normal distribution.

$$\ln S_H \sim N\big(\ln S_0 + (\mu - \frac{1}{2}diagonal(\Sigma))H, \Sigma H\big),$$

where $diagonal(\Sigma) = \{\Sigma_{11}, \ldots, \Sigma_{DD}\}^T$ is the diagonal vector.

$$S_T^d = S_H^d \exp\{(r - \frac{1}{2}\Sigma_{dd})(T - H) + \sum_{j=1}^{m} L_{dj}(W_T^j - W_H^j)\}, \quad d = 1, \ldots, D$$

$$\ln \frac{S_T^d}{S_H^d} = (r - \frac{1}{2}\Sigma_{dd})(T - H) + \sum_{j=1}^{m} L_{dj}(W_T^j - W_H^j), \quad d = 1, \ldots, D$$

and

$$\ln \frac{S_T^d}{S_H^d} \sim N\big((r - \frac{1}{2}\Sigma_{dd})(T - H), \sum_{j=1}^{m} L_{dj}^2(T - H)\big), \quad d = 1, \ldots, D$$

The D-dimensional vector $\ln \frac{S_T}{S_H} \triangleq \{\ln(\frac{S_T^1}{S_H^1}), \ldots, \ln(\frac{S_T^D}{S_H^D})\}$ follows multivariate normal distribution.

$$\ln \frac{S_T}{S_H} \sim N\big((r - \frac{1}{2}diagonal(\Sigma))(T - H), \Sigma(T - H)\big).$$

Then $\ln S_T = \ln(\frac{S_T}{S_H}) + \ln S_H$, is the sum of two independent multi-normal variable, which still follows multivariate normal distribution, with mean $\bar{S}_1 \triangleq \ln S_0 + (\mu - \frac{1}{2}diagonal(\Sigma))H + (r - \frac{1}{2}diagonal(\Sigma))(T - H)$ and covariance matrix ΣT. Hence the density function of $\ln S_T$ is

$$f_1(u) = \frac{1}{(2\pi)^{D/2}det(\Sigma T)^{1/2}} \exp\{-\frac{1}{2}(u - \bar{S}_1)^T(\Sigma T)^{-1}(u - \bar{S}_1)\}.$$

The density function of S_T is $F(u) = f_1(\ln u)/u_1 u_2 \ldots u_D$ where u is a D-dimensional vector and u_i is the ith element of vector u.

A.2. Proof of Theorem 3.2

Under risk-neutral measure, given that $S_H^d = x^d$, S_T^d is simulated by

$$\begin{aligned} S_T^d &= S_H^d \cdot \exp[(r - \frac{1}{2}\Sigma_{dd})(T-H) + \sum_{j=1}^{m} L_{dj}(W_T^j - W_H^j)] \\ &= x^d \cdot \exp[(r - \frac{1}{2}\Sigma_{dd})(T-H) + \sum_{j=1}^{m} L_{dj}(W_T^j - W_H^j)] \quad for\ d = 1, 2, \ldots, D. \end{aligned}$$

Thus, the D-dimensional vector $\ln S_T$ follows multivariate normal distribution.

$$\ln S_T \sim N\big(\ln x + (r - \frac{1}{2}diagonal(\Sigma))(T-H), \Sigma(T-H)\big),$$

where $diagonal(\Sigma) = \{\Sigma_{11}, \ldots, \Sigma_{DD}\}^T$ is the diagonal vector.

Given $\ln S_H = \ln x$, the density function for $\ln S_T$ is

$$f_2(x,u) = \frac{1}{(2\pi)^{D/2} det(\Sigma(T-H))^{1/2}} \exp\{-\frac{1}{2}(u - \bar{S}_2)^T (\Sigma(T-H))^{-1}(u - \bar{S}_2)\}.$$

where $\bar{S}_2 \triangleq \ln x + (r - \frac{1}{2}diagonal(\Sigma))(T-H)$.

Thus the transition density of S_T given $S_H = x$ is $f_T(x,u) = f_2(x, \ln u)/u_1 u_2 \ldots u_D$ where u is a D-dimensional vector and u_i is the ith element of vector u.

Bibliography

[1] Acerbi, C., Tasche, D.(2002) *On the coherence of expected shortfall.* J. Banking Finance 26, 1487-1503

[2] Acerbi, C., Tasche, D. (2002) *Expected Shortfall: a natural coherent alternative to Value at Risk.* Economic Notes 31(2), 379-388.

[3] Artzner, P., Delbaen, F., Eber, J-M., Heath, D. (1997) *Thinking coherently,* Risk, 10, 68-71.

[4] Artzner, P., Delbaen, F., Eber, J-M., Heath, D. (1999) *Coherent measures of risk,* Mathematical Finance, 9, 3, 203-228.

[5] Broadie, M, Y. Du, C. C. Moallemi. (2011) *Efficient risk estimation via nested sequential simulation.* Management Science, 57(6):1172-1194 .

[6] Broadie,M, P. Glasserman (2004). *A stochastic mesh method for pricing high-dimensional American options.* Journal of Computational Finance, 7(4), 35-72.

[7] Crouhy, M, D. Galai, and R. Mark. (2000) *Risk Management.* McGraw-Hill, New York,

[8] Duffie, D. and J. Pan, (1997), *An overview of value at risk* Journal of Derivatives, Vol. 4, No. 3 , pp. 7-49

[9] Embrechts, P. (2000) *Extreme Value Theory: Potential and Limitations as an Integrated Risk Management Tool.* Derivatives Use, Trading & Regulation 6, 449-456

[10] Frye, J. (1998). *Monte Carlo by day.* Risk 11:66-71.

[11] Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering.* Springer-Verlag, New York.

[12] Gordy, M. B., and S. Juneja. (2006). *Efficient simulation for risk measurement in portfolio of CDOs.* In Proceeedings of the 2006 Winter Simulation Conference, ed. L. F. Perrone, B. Lawson, J. Liu, and F. P. Wieland, 749-756: IEEE Press.

[13] Gordy, M. B., and S. Juneja. (2008). *Nested simulation in portfolio risk measurement.* Finance and Economics Discussion Series 2008-21, Federal Reserve Board.

[14] Jorion, P., (1997). *Value at Risk.* McGraw-Hill, New York

[15] Lan, H., B. L. Nelson, and J. Staum. (2007). *A confidence interval for tail conditional expectation via two-level simulation.* In Proceedings of the 2007 Winter Simulation Conference, ed. S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 949-957: IEEE Press.

[16] Lan, H., B. L. Nelson, and J. Staum. (2010). *A confidence interval procedure for expected shortfall via two-level simulation.* Operations Research. In press.

[17] Lee, S.-H. (1998). *Monte Carlo computation of conditional expectation quantiles.* Ph. D. thesis, Stanford University

[18] Linsmeier, T. J., N. D. Pearson. (1996). *Risk Measurement: An introduction to value-at-risk Technical report* 9 6-04, OFOR, University of Illinois, Urbana-Champaign, IL

[19] Liu, M, Jeremy Staum. (2009) *Estimating Expected Shortfall with Stochastic Kriging.* Winter Simulation Conference' 2009. pp.1249 1260

[20] Longstaff, F., E. S. Schwartz. (2001). *Valuing American options by simulation: A simple least squares approach.* Rev. Financial Stud. 14 113-147.

[21] McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton University Press, Princeton

[22] Rootzen, H., Kluppelberg, C., (1999). *A single number cannot hedge against economic catastrophes.* Ambio-Royal Swedish Academy of Sciences 28, 550-555.

[23] Shaw, J. (1998). *Beyond VAR and stress testing. In Monte Carlo: Methodologies and Applications for Pricing and Risk Management*, ed. B. Dupire, 231-244. London: Risk Books.

[24] Yamai, Y., and T. Yoshiba, (2002c) *Comparative Analyses of Expected Shortfall and VaR (2): Expected Utility Maximization and Tail Risk* Monetary and Economic Studies, Vol. 20, No. 2, Bank of Japan, forthcoming,